





Le Frido 2025  
volume 3  
Laurent Claessens

Plusieurs extensions et versions de ce livre.

1. La version courante, régulièrement mise à jour et qui deviendra petit à petit le Frido 2026.  
Téléchargeable sur

<https://laurent.claessens-donadello.eu/pdf/lefrido.pdf>

2. La version la plus complète, contenant beaucoup de géométrie différentielle

<https://laurent.claessens-donadello.eu/pdf/giulietta.pdf>

3. Et bien entendu les sources  $\text{\LaTeX}$

<https://github.com/LaurentClaessens/mazhe>

Copyright 2011-2025 Laurent Claessens, and many contributors. A complete list could be retrieved from the git log.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the chapter entitled “GNU Free Documentation License”.

(c) 2015-2022 David Revoy pour les illustrations de couverture CC-BY,  
<https://www.peppercarrot.com/>

ISBN : 979-10-97085-44-5

# Chapitre 16

## Représentations et caractères

### 16.1 Représentations et caractères

#### Définition 16.1.

Si  $G$  est un groupe, l'ensemble des homomorphismes  $\text{Hom}(G, \mathbb{C}^*)$  est un groupe pour la multiplication. Un élément de  $\text{Hom}(G, \mathbb{C}^*)$  est un **caractère abélien**. Le nom « abélien » vient du fait que le caractère prenne ses valeurs dans  $\mathbb{C}^*$ . Nous notons  $\hat{G} = \text{Hom}(G, \mathbb{C}^*)$ .

#### Théorème 16.2.

Soit  $G$  un groupe abélien fini. Alors  $G$  est isomorphe à  $\hat{\hat{G}}$ .

L'isomorphisme n'est pas canonique.

*Démonstration.* Étant donné la structure des groupes abéliens finis donnée par le théorème 5.23, nous commençons par nous concentrer sur  $G = \mathbb{Z}/n\mathbb{Z}$ . Nous allons montrer que

$$\text{Hom}(\mathbb{Z}/n\mathbb{Z}) \simeq U_n = \{\xi \in \mathbb{C} \text{ tel que } \xi^n = 1\}. \quad (16.1)$$

Pour cela nous avons l'isomorphisme

$$\begin{aligned} \psi: \text{Hom}(\mathbb{Z}, \mathbb{C}^*) &\rightarrow \mathbb{C}^* \\ f &\mapsto f(1). \end{aligned} \quad (16.2)$$

Notons que si  $f \in \text{Hom}(\mathbb{Z}, \mathbb{C}^*)$ , alors  $f(k) = f(1)^k$ , donc  $\psi$  est bien un isomorphisme. Cela nous amène à définir

$$\begin{aligned} \varphi: \text{Hom}\left((\mathbb{Z}/n\mathbb{Z}, +), (\mathbb{C}^*, \cdot)\right) &\rightarrow U_n \\ f &\mapsto f(1). \end{aligned} \quad (16.3)$$

Remarquons que pour tout  $f \in \text{Hom}(\mathbb{Z}/n\mathbb{Z}, \mathbb{C}^*)$  on a bien  $f(1)^n = 1$ . En effet si  $[k] \in \mathbb{Z}/n\mathbb{Z}$ , alors  $f([k]) = f(1)^k$  et en particulier

$$f(1)^n = f([n]) = f(0) = 1. \quad (16.4)$$

Donc  $f(1) \in U_n$ . Le  $\varphi$  est injective parce que si  $f(1) = g(1)$  alors  $f = g$  du fait que  $f(k) = f(1)^k = g(1)^k = g(k)$ .

Nous en sommes à avoir prouvé que  $\text{Hom}(\mathbb{Z}/n\mathbb{Z}, \mathbb{C}^*) \simeq U_n$  (introduit au lemme 19.2). Il faudrait encore montrer que  $U_n \simeq \mathbb{Z}/n\mathbb{Z}$ . Pour cela nous nous rappelons du lemme 19.7 nous ayant raconté que le groupe  $U_n$  des racines de l'unité était cyclique et d'ordre  $n$ . Il est donc bien isomorphe à  $\mathbb{Z}/n\mathbb{Z}$ .

Passons au cas où

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z} \times \dots \times \mathbb{Z}/d_k\mathbb{Z}. \quad (16.5)$$

Dans ce cas nous montrons que

$$\begin{aligned} \alpha: \bigtimes_{i=1}^k \operatorname{Hom}(\mathbb{Z}/d_i\mathbb{Z}, \mathbb{C}^*) &\rightarrow \operatorname{Hom}(G, \mathbb{C}^*) \\ \alpha(\chi_1, \dots, \chi_k)(g_1, \dots, g_k) &= \chi_1(g_1) \dots \chi_k(g_k). \end{aligned} \quad (16.6)$$

Ce  $\alpha$  est injectif parce qu'en appliquant l'égalité

$$\alpha(\chi_1, \dots, \chi_k) = \alpha(\chi'_1, \dots, \chi'_k) \quad (16.7)$$

à l'élément  $g = (0, \dots, 1, \dots, 0)$  alors nous trouvons  $\chi_i(1) = \chi'_i(1)$  parce que  $\chi_j(0) = 1$ . Du coup  $\chi_i = \chi'_i$ .

L'application  $\alpha$  est en plus surjective. En effet si  $\chi \in \operatorname{Hom}(G, \mathbb{C}^*)$ , alors nous définissons

$$\chi_i(g_i) = \chi(0, \dots, g_i, \dots, 0), \quad (16.8)$$

et nous avons alors  $\alpha(\chi_1, \dots, \chi_k) = \chi$ .

Nous devons encore montrer que  $\alpha$  est un homomorphisme. Si  $\chi, \chi' \in \bigtimes_{i=1}^k \operatorname{Hom}(\mathbb{Z}/d_i, \mathbb{C}^*)$ , alors

$$\alpha(\chi\chi')(g_1, \dots, g_k) = (\chi_1\chi'_1)(g_1) \dots (\chi_k\chi'_k)(g_k) \quad (16.9a)$$

$$= \chi_1(g_1) \dots \chi_k(g_k) \chi'_1(g_1) \dots \chi'_k(g_k) \quad (16.9b)$$

$$= \alpha(\chi)(g_1, \dots, g_k) \alpha(\chi')(g_1, \dots, g_k) \quad (16.9c)$$

$$= (\alpha(\chi)\alpha(\chi'))(g_1, \dots, g_k). \quad (16.9d)$$

Donc  $\alpha(\chi\chi') = \alpha(\chi)\alpha(\chi')$ . □

### Théorème 16.3.

Soit  $G$  un groupe abélien fini. Les groupes  $G$  et  $\hat{\hat{G}}$  sont isomorphes et un isomorphisme canonique est donné par  $\alpha: g \mapsto f_g$  donné par

$$f_g(\chi) = \chi(g). \quad (16.10)$$

*Démonstration.* D'abord  $f_g$  est bien un caractère de  $\hat{G}$  parce que

$$f_g(\chi\chi') = (\chi\chi')(g) = \chi(g)\chi'(g) = f_g(\chi)f_g(\chi'). \quad (16.11)$$

Le fait que  $\alpha$  soit un homomorphisme de groupes est direct :

$$f_{gg'}(\chi) = \chi(gg') = \chi(g)\chi(g') = f_g(\chi)f_{g'}(\chi) = (f_g f_{g'}) (\chi). \quad (16.12)$$

D'autre part nous savons que  $G$  et  $\hat{\hat{G}}$  ont le même cardinal. Il suffit donc de prouver l'injectivité de  $\alpha$  pour être sûr de la bijectivité. Pour cela nous devons prouver que si  $g \neq e$  alors  $f_g \neq f_e$ . Nous savons que pour tout caractère  $\chi \in \hat{G}$ ,  $f_e(\chi) = \chi(e) = 1$ . Donc pour tout  $g \in G \setminus \{e\}$ , nous devons trouver  $\chi \in \hat{G}$  tel que  $\chi(g) \neq 1$ .

En vertu de ce que nous connaissons sur la structure des groupes abéliens finis (théorème 5.23), nous commençons  $G = \mathbb{Z}/n\mathbb{Z}$  et considérons le caractère donné par  $\chi([1]) = e^{2i\pi/n}$ . Ce  $\chi$  est un isomorphisme entre  $G$  et  $U(n)$ ; nous n'avons  $\chi([k]) = 0$  que si  $[k] = [n] = [0]$ . Pour rappel dans  $\mathbb{Z}/n\mathbb{Z}$ , le neutre est  $e = 0$  et non  $e = 1$ .

Passons au cas général :

$$G \simeq \mathbb{Z}/n_1\mathbb{Z} \times \dots \times \mathbb{Z}/n_k\mathbb{Z} \quad (16.13)$$

Si  $g = (g_1, \dots, g_k)$  est non nul dans  $G$ , alors il existe  $k$  tel que  $g_k \neq 0$  et on prend

$$\chi(g_1, \dots, g_k) = \chi_k(g_k) \quad (16.14)$$

où  $\chi_k$  est le caractère  $\chi_i([1]) = e^{2\pi i/n_k}$ . Ce  $\chi$  est alors un caractère non trivial de  $G$ . □

### 16.1.1 Crochet de dualité et transformée de Fourier

Si  $G$  est un groupe abélien, nous définissons le crochet de dualité entre  $G$  et  $\hat{G}$  par

$$\begin{aligned}\langle \cdot, \cdot \rangle: G \times \hat{G} &\rightarrow \mathbb{C}^* \\ \langle g, \chi \rangle &= \chi(g).\end{aligned}\tag{16.15}$$

Notons que l'image de ce crochet n'est pas  $\mathbb{C}^*$  entier, mais seulement le groupe unitaire  $U(n)$  où  $n$  est l'exposant <sup>1</sup> de  $G$ .

Si  $f, g$  sont des applications de  $G$  dans  $\mathbb{C}$ , alors on leur associe le produit scalaire

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{s \in G} f(s) \overline{g(s)}.\tag{16.16}$$

#### Lemme 16.4.

Les caractères de  $G$  forment une base orthonormée de  $\mathbb{C}^G$  pour ce produit scalaire.

*Démonstration.* Étant donné que les  $\chi(s)$  sont des nombres complexes de module 1, nous avons  $\chi(s)\overline{\chi(s)} = 1$  et par conséquent  $\langle \chi, \chi \rangle = 1$ .

Si par contre  $\chi \neq \chi'$ , alors il existe  $s_0 \in G$  tel que  $\chi(s_0) \neq \chi'(s_0)$ . Dans ce cas en effectuant un changement de variable  $s \rightarrow s_0 s$  dans la sommation,

$$\langle \chi, \chi' \rangle = \frac{1}{|G|} \sum_{s \in G} \chi(s) \overline{\chi'(s)}\tag{16.17a}$$

$$= \frac{1}{|G|} \sum_{s \in G} \chi(s_0 s) \overline{\chi'(s_0 s)}\tag{16.17b}$$

$$= \frac{1}{|G|} \chi(s_0) \overline{\chi'(s_0)} \sum_{s \in G} \chi(s) \overline{\chi'(s)}.\tag{16.17c}$$

Donc nous avons trouvé

$$\langle \chi, \chi' \rangle (1 - \chi(s_0) \overline{\chi'(s_0)}) = 0.\tag{16.18}$$

Mais vu que  $\chi(s_0) \neq \chi'(s_0)$ , la parenthèse est non nulle (pour rappel  $\chi(s_0)$  est un complexe de module 1) et par conséquent  $\langle \chi, \chi' \rangle = 0$ .

Nous déduisons immédiatement que les caractères forment une famille libre parce que si  $\sum_i a_i \chi_i = 0$  (la somme est sur tous les caractères), alors en prenant le produit scalaire avec  $\chi_k$ ,

$$\sum_i a_i \langle \chi_k, \chi_i \rangle = 0,\tag{16.19}$$

et donc  $a_k = 0$ .

Les caractères forment donc un système libre orthonormé. De plus l'espace engendré à la bonne dimension parce que le cardinal de l'ensemble des caractères est la dimension (complexe) de l'espace des fonctions de  $G$  dans  $\mathbb{C}$  parce que, en utilisant l'isomorphisme entre  $G$  et  $\hat{G}$ ,

$$\text{Card } \hat{G} = \text{Card}(G) = \dim_{\mathbb{C}} \mathbb{C}^G.\tag{16.20}$$

□

Du fait que les caractères forment une base orthonormée, nous pouvons écrire, pour toute application  $f: G \rightarrow \mathbb{C}$ ,

$$f = \sum_{\chi \in \hat{G}} \langle \chi, f \rangle \chi.\tag{16.21}$$

À une fonction  $f: G \rightarrow \mathbb{C}$  nous associons la **transformée de Fourier**

$$\begin{aligned}\hat{f}: \hat{G} &\rightarrow \mathbb{C} \\ \chi &\mapsto \langle \chi, f \rangle.\end{aligned}\tag{16.22}$$

1. Définition 1.413.



Nous avons donc aussi une espèce de formule d'inversion

$$f = \sum_{\chi \in \hat{G}} \hat{f}(\chi) \chi \quad (16.23)$$

qui n'est qu'une réécriture de 16.21.

### 16.1.2 Groupes non abéliens

Nous avons vu que le groupe des caractères  $\hat{G}$  contenait toute l'information sur un groupe abélien. Malheureusement, pour les groupes non abéliens, ça ne va pas suffire, et nous allons introduire la notion de représentations, dont les caractères seront un cas particulier de dimension un.

#### Proposition 16.5.

Soit  $G$  un groupe (pas spécialement abélien). Nous avons

$$\hat{G} \simeq \text{Hom}(G/D(G), \mathbb{C}^*). \quad (16.24)$$

*Démonstration.* Ce qui fait fonctionner la preuve est le fait que si  $f: G \rightarrow \mathbb{C}^*$  est un homomorphisme, alors  $f$  s'annule sur  $D(G)$ . L'isomorphisme est

$$\begin{aligned} \psi: \hat{G} &\rightarrow \text{Hom}(G/D(G), \mathbb{C}^*) \\ \psi(f)[g] &= f(g). \end{aligned} \quad (16.25)$$

Cette application est bien définie parce que si  $f$  est un homomorphisme,

$$f(gklk^{-1}l^{-1}) = f(g). \quad (16.26)$$

D'autre part  $\psi$  est un homomorphisme de groupe parce que

$$\psi(f_1 f_2)[g] = (f_1 f_2)(g) = f_1(g) f_2(g) = \psi(f_1)[g] \psi(f_2)[g] = (\psi(f_1) \psi(f_2))[g]. \quad (16.27)$$

Pour l'injectivité de  $\psi$ , soit  $f_1$  et  $f_2$  telles que  $\psi(f_1) = \psi(f_2)$ . Alors pour tout  $g \in G$  nous avons

$$\psi(f_1)[g] = \psi(f_2)[g] \quad (16.28)$$

et donc  $f_1(g) = f_2(g)$ .

Enfin  $\psi$  est surjective. En effet, soit  $\bar{f} \in \text{Hom}(G/D(G), \mathbb{C}^*)$ . Alors nous obtenons  $\psi(f) = \bar{f}$  en posant

$$f(g) = \bar{f}[g]. \quad (16.29)$$

Il faut juste vérifier que le  $f$  ainsi défini est dans  $\hat{G}$ , c'est-à-dire que  $f(g_1 g_2) = f(g_1) f(g_2)$ .  $\square$

Cette proposition nous montre que

$$\hat{G} = \widehat{G/D(G)}, \quad (16.30)$$

alors que  $G/D(G)$  est abélien; il n'est donc pas tellement possible que  $\hat{G}$  contienne beaucoup d'informations intéressantes sur  $G$ .

### 16.1.3 Représentations linéaires des groupes finis

Si  $\dim V = 1$ , alors  $\text{GL}(V) = \mathbb{C}^*$  et les représentation sont les caractères abéliens.

ExKUAYUD

**Exemple 16.6.**

Considérons le triangle équilatéral  $A, B, C$  donné par les points

$$\left\{ \begin{array}{l} A = 1 \end{array} \right. \quad (16.31a)$$

$$\left\{ \begin{array}{l} B = (-\frac{1}{2}, \frac{\sqrt{3}}{2}) \end{array} \right. \quad (16.31b)$$

$$\left\{ \begin{array}{l} C = (-\frac{1}{2}, -\frac{\sqrt{3}}{2}) \end{array} \right. \quad (16.31c)$$

$$\quad (16.31d)$$

Dans la base (pas orthonormée)  $\{A, B\}$  de  $\mathbb{R}^2$ , ces trois points sont donnés par

$$A = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad C = \begin{pmatrix} -1 \\ -1 \end{pmatrix}. \quad (16.32)$$

Le groupe symétrique  $S_3$  agit sur le triangle par permutation des sommets. Vues dans la base  $\{A, B\}$ , les transpositions correspondent aux matrices

$$\rho(AB) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (16.33a)$$

$$\rho(AC) = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \quad (16.33b)$$

$$\rho(BC) = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}. \quad (16.33c)$$

Cherchons la matrice correspondante à la permutation  $(A, B, C)$ . Vu que  $A$  est envoyé sur  $B$ , la première colonne sera <sup>2</sup>  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , et comme  $B$  est envoyé sur  $C$ , la deuxième colonne sera  $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ . Nous avons donc

$$\rho(ABC) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}. \quad (16.34)$$

Nous vérifions que  $\rho(ABC)C = A$  :

$$\rho(ABC)C = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = A \quad (16.35)$$

Par ailleurs, la permutation  $(A, B, C)$  se décompose en  $(A, B, C) = (A, C)(A, B)$  et nous pouvons vérifier que

$$\rho(AC)\rho(AB) = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} = \rho(ABC). \quad (16.36)$$

△

**Définition 16.7.**

Si  $(V, \rho)$  et  $(V', \rho')$  sont deux représentations du groupe  $G$ , alors nous définissons la **somme directe** par  $(V \oplus V', \rho \oplus \rho')$  donné par

$$(\rho \oplus \rho')(g) = \begin{pmatrix} \rho(g) & 0 \\ 0 & \rho'(g) \end{pmatrix} \in \text{GL}(V \oplus V'). \quad (16.37)$$

Nous noterons souvent  $2V$  pour la représentations  $(V, \rho) \oplus (V, \rho)$  et plus généralement l'écriture

$$V = \bigoplus_i k_i W_i \quad (16.38)$$

signifiera la représentation somme de  $k_i$  termes de la représentation  $W_i$ . Ici encore un abus est commis entre la représentation  $(\rho_i, W_i)$  et l'espace  $W_i$ .

2. Les colonnes sont les images des vecteurs de base.

### 16.1.4 Module

Nous considérons la  $\mathbb{C}$ -algèbre  $G[\mathbb{C}]$  des combinaisons (formelles) d'éléments de  $G$  à coefficients dans  $\mathbb{C}$ , c'est-à-dire l'ensemble

$$\mathbb{C}[G] = \left\{ \sum_{s \in G} a_s s \right\} \quad (16.39)$$

avec le produit hérité de la bilinéarité :

$$\sum_{s \in G} \sum_{t \in G} a_s b_t st = \sum_s \sum_t a_s b_{s^{-1}t} t, \quad (16.40)$$

et la somme

$$\left( \sum_s a_s s \right) + \sum_t b_t t = \sum_{s \in G} (a_s + b_s) s. \quad (16.41)$$

Le tout est une  $\mathbb{C}$ -algèbre agissant sur  $V$  par

$$\left( \sum_s a_s s \right) v = \sum_{s \in G} a_s \rho(s) v \in V \quad (16.42)$$

Les sous-modules indécomposables seront les représentations irréductibles.

#### Définition 16.8.

La représentation  $(V, \rho)$  du groupe  $G$  est **irréductible** si les seuls sous-espaces invariants de  $V$  sous  $\rho(G)$  sont  $V$  et  $\{0\}$ .

#### Exemple 16.9.

La représentation de  $S_3$  sur  $\mathbb{R}^2$  donnée par les permutations des sommets d'un triangle équilatéral donnée dans l'exemple 16.6 est irréductible.  $\triangle$

La question qui vient est de savoir si une représentation possédant des sous-espaces invariants peut être écrite comme la somme de représentations irréductibles.

PropHeyoAN

#### Proposition 16.10.

Soit  $(V, \rho)$  une représentation linéaire de dimension finie d'un groupe fini<sup>3</sup>. Si  $W_1$  est un sous-espace stable<sup>4</sup>, alors il existe un sous-espace  $W_2$  également stable et tel que  $V = W_1 \oplus W_2$ .

Toute représentation linéaire est décomposable en représentations irréductibles.

*Démonstration.* Soit  $P: V \rightarrow V$  un projecteur sur  $W_1$ , c'est-à-dire que  $P^2 = P$  et  $P(V) = W_1$ . Pour construire un tel projecteur, on peut par exemple prendre un supplémentaire de  $W_1$  dans  $V$  puis utiliser la décomposition<sup>5</sup>. Nous considérons l'opérateur

$$P_G = \frac{1}{|G|} \sum_{g \in G} \rho(g) \circ P \circ \rho(g)^{-1}. \quad (16.43)$$

Prouvons que ce  $P_G$  est encore un projecteur. D'abord pour tout  $g \in G$  nous avons

$$\rho(g) P_G \rho(g)^{-1} = \frac{1}{|G|} \sum_{s \in G} \rho(gs) P \rho(gs)^{-1} = P_G. \quad (16.44)$$

La dernière égalité est un changement de variables dans la somme<sup>6</sup>. Cela signifie que  $P_G \rho = \rho P_G$ .

3. La démonstration marche aussi pour les groupes compacts, mais il faudrait des intégrales.

4. c'est-à-dire si  $\rho$  n'est pas irréductible.

5. Ou encore prendre une base de  $W_1$ , l'étendre en une base de  $V$  et définir  $P$  comme l'annulation des coefficients des vecteurs « complétant » la base.

6. Et c'est ça qui demande un peu de technique pour écrire la preuve dans le cas d'un groupe compact : il faut une mesure de Haar.

Nous avons même  $P_G P = P$  parce que si  $v \in W_1$ , alors

$$P_G(v) = \frac{1}{|G|} \sum_{s \in G} \rho(s) P \underbrace{\rho(s)^{-1} v}_{\in W_1} \quad (16.45a)$$

$$= \frac{1}{|G|} \sum_s \rho(s) \rho(s)^{-1} v \quad (16.45b)$$

$$= v. \quad (16.45c)$$

Avec cela nous pouvons conclure que  $P_G^2 = P_G$  parce que

$$P_G \circ P_G = \frac{1}{|G|} \sum_g P_G \rho(g) P \rho(g)^{-1} \quad (16.46a)$$

$$= \frac{1}{|G|} \sum_g \rho(g) P_G P \rho(g)^{-1} \quad (16.46b)$$

$$= \frac{1}{|G|} \sum_g \rho(g) P \rho(g)^{-1} \quad (16.46c)$$

$$= P_G. \quad (16.46d)$$

Donc  $P_G$  est un projecteur, est stable sous les conjugaisons par  $\rho(g)$  et commute avec  $\rho(g)$ . Nous décomposant  $\text{Id}$  de façon évidente en

$$\text{Id} = P_G + (\text{Id} - P_G). \quad (16.47)$$

Étant donné que l'opérateur  $P_G$  commute avec tous les  $\rho(g)$ , les noyaux de  $P_G$  et  $\text{Id} - P_G$  sont des sous-espaces invariants. Vu que  $P_G$  est un projecteur, nous avons  $q(P_G) = 0$  avec  $q(X) = X^2 - X$ . Pour appliquer le lemme des noyaux (théorème 8.91), nous remarquons que  $q(X) = X(X - 1)$  et donc

$$V = \ker P_G \oplus \ker(P_G - \mathbb{1}). \quad (16.48)$$

Si nous posons  $W_2 = \ker P_G$ , il reste à voir que  $\ker(P_G - \mathbb{1}) = W_1$ . D'abord  $W_1 \subset \ker(P_G - \text{Id})$  parce que si  $w \in W_1$ , ce dernier étant stable,

$$P_G w = \frac{1}{|G|} \sum_{g \in G} \rho(g) P \underbrace{\rho(g)^{-1} w}_{\in W_1} \quad (16.49a)$$

$$= \frac{1}{|G|} \sum_{g \in G} w \quad (16.49b)$$

$$= w. \quad (16.49c)$$

Pour prouver l'inclusion inverse, nous savons que  $P_G$  et  $P$  sont des projecteurs tels que  $P_G P = P$ , ce qui signifie que l'image de  $P_G$  est incluse à celle de  $P$ , c'est-à-dire à  $W_1$ . Mais  $\text{Image}(P_G) = \ker(\mathbb{1} - P_G)$ , donc

$$\ker(\mathbb{1} - P_G) = \text{Image}(P_G) \subset \text{Image}(P) = W_1. \quad (16.50)$$

La représentation  $\rho$  se décompose donc en deux sous-représentations  $(\rho, W_1)$  et  $\rho, W_2$ . Si l'une des deux n'est pas irréductible, le processus peut recommencer. Vu que la dimension de  $V$  est finie, toute représentation se décompose en une somme finie de représentation irréductibles.  $\square$

### 16.1.5 Structure hermitienne

Soit  $(\rho, V)$  une représentation de  $G$  sur un espace vectoriel complexe  $V$ . Nous voulons munir  $V$  d'un produit scalaire hermitien (définition 8.172) tel que les opérateurs  $\rho(g)$  soient tous des isométries. C'est-à-dire que nous voudrions définir  $\langle u, v \rangle_G$  de telle sorte à avoir

$$\langle \rho(g)u, \rho(g)v \rangle_G = \langle u, v \rangle_G \quad (16.51)$$

pour tout  $g \in G$ . Nous commençons par considérer un produit hermitien  $\langle \cdot, \cdot \rangle$  quelconque et puis nous définissons

$$\langle u, v \rangle_G = \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)u, \rho(g)v \rangle. \quad (16.52)$$

Nous devons vérifier que c'est un produit. La seule des conditions dont la vérification n'est pas immédiate est celle de positivité. Pour tout  $g \in G$  et tout  $v \in V$ , nous avons  $\langle \rho(g)v, \rho(g)v \rangle$  est positif et nul si et seulement si  $\rho(g)v = 0$ . Étant donné que  $\rho(e)v = v$ , parmi les termes de la somme

$$\langle u, u \rangle_G = \frac{1}{|G|} \sum_{g \in G} \langle \rho(g)v, \rho(g)v \rangle, \quad (16.53)$$

au moins un est strictement positif (pourvu que  $v \neq 0$ ); les autres sont positifs ou nuls. Par conséquent  $\langle v, v \rangle_G = 0$  si et seulement si  $v = 0$ .

Donc les groupes finis peuvent être vus comme des parties de groupes d'isométrie. De la même façon, en utilisant une mesure de Haar pour faire la moyenne, nous pouvons plonger les groupes compacts dans des groupes unitaires.

### 16.1.6 Caractères

#### Définition 16.11.

Soit  $(V, \rho)$  une représentation linéaire du groupe  $G$ . Le **caractère** de  $\rho$  est la fonction

$$\begin{aligned} \chi_\rho: G &\rightarrow \mathbb{C} \\ s &\mapsto \text{Tr}(\rho(s)). \end{aligned} \quad (16.54)$$

Par invariance cyclique de la trace, nous avons

$$\chi_\rho(sts^{-1}) = \chi_\rho(t), \quad (16.55)$$

ce qui fait que le caractère est une fonction constante sur les classes de conjugaison.

D'après sa fiche wikipédia, le marquis de Sade, passionné de théâtre, faisait des représentations qui avaient du caractère.

Un **caractère irréductible** est un caractère d'une représentation irréductible.

#### Définition 16.12.

Une application  $f: G \rightarrow \mathbb{C}$  est **centrale** si elle est constante sur les classes de conjugaison.

Les traces sont des applications centrales.

L'ensemble des fonctions centrales sur un groupe fini (ou tout au moins ayant un nombre fini de classes de conjugaison) est un  $\mathbb{C}$ -espace vectoriel de dimension égale au nombre de classes, et nous pouvons mettre le produit scalaire

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{s \in G} f(s) \overline{g(s)}. \quad (16.56) \quad \text{EqJrEpVI}$$

C'est une forme hermitienne sur l'espace des fonctions centrales.

## 16.2 Équivalence de représentations et caractères

Cette section prend des éléments des articles **lemme de Schur**, **caractère d'une représentation**, **fonction centrale** et **trace** de wikipédia.

#### Définition 16.13.

Nous disons que les deux représentations  $(V, \rho)$  et  $(V', \rho')$  sont **équivalentes** si il existe une bijection linéaire  $f: V \rightarrow V'$  telle que

$$f \circ \rho = \rho' \circ f. \quad (16.57)$$

Nous disons alors que  $f$  **entrelace**  $\rho$  et  $\rho'$ .

ThoyftobH

**Théorème 16.14** (Théorème de Schur).

Si  $(V, \rho)$  et  $(V', \rho')$  sont des représentations irréductibles non équivalentes alors la seule application linéaire  $f: V \rightarrow V'$  entretenant  $\rho$  et  $\rho'$  est la fonction nulle.

En d'autres termes, soit les représentations sont équivalentes (et il y a un isomorphisme), soit il n'y a même pas un homomorphisme.

*Démonstration.* Soit  $f \in \mathcal{L}(V, V')$  telle que  $f \circ \rho = \rho' \circ f$ . Alors  $\ker f$  est un sous-espace stable sous  $\rho(G)$ , et  $\text{Image}(f)$  est un sous-espace de  $V'$  stable par  $\rho'(G)$ . Par irréductibilité, nous avons que  $\ker(f) = \{0\}$  ou  $V$ . Même chose pour  $\text{Image}(f)$ . Il y a deux possibilités.

- (1) Si  $\ker(f) = \{0\}$ , alors  $\text{Image}(f) \neq \{0\}$  et alors  $\text{Image}(f) = V'$ . Du coup  $f$  est injective et surjective, c'est-à-dire est un isomorphisme.
- (2) Si  $\ker(f) = V$ , alors  $f = 0$ .

□

**Corolaire 16.15** (Schur pour les représentations sur  $\mathbb{C}$ ).

Soit  $(V, \rho)$  une représentation irréductible, alors l'ensemble

$$\text{End}_G(V, \rho) = \{f \in \text{End}(V) \text{ tel que } \rho \circ f = f \circ \rho\} \quad (16.58)$$

est l'ensemble des homothéties.

*Démonstration.* Soit  $f \in \text{End}_G(V, \rho)$ . Vu que l'espace est sur  $\mathbb{C}$ , l'endomorphisme  $f$  a une valeur propre  $\lambda$ . L'opérateur  $g = f - \lambda \mathbb{1}$  est aussi un opérateur d'entrelacement de  $\rho$  alors que  $\ker(g) \neq \{0\}$  par définition de valeur propre. Du coup  $\ker(g) = V$ , ce qui signifie que  $f$  est l'isométrie de rapport  $\lambda : f = \lambda \text{Id}$ .

□

LempUSOLO

**Lemme 16.16.**

Si  $(\rho, V)$  et  $(\rho', V')$  sont des représentations équivalentes de caractères  $\chi$  et  $\chi'$ , alors  $\chi = \chi'$ .

*Démonstration.* Si  $A: V \rightarrow V'$  est un isomorphisme d'espace vectoriel entretenant  $\rho$  et  $\rho'$ , c'est-à-dire si pour tout  $g$ ,  $\rho'(g)A = A\rho(g)$ , alors  $\rho'(g) = A\rho(g)A^{-1}$  et

$$\chi'(g) = \text{Tr}(\rho'(g)) = \text{Tr}(A\rho(g)A^{-1}) = \text{Tr}(\rho(g)) \quad (16.59)$$

parce que la trace est un invariant de similitude (lemme 8.207).

□

LempJqIZns

**Lemme 16.17.**

Si  $\chi$  est le caractère de la représentation complexe  $(V, \rho)$  du groupe fini  $G$ , alors pour tout  $g \in G$  nous avons  $\chi(g^{-1}) = \overline{\chi(g)}$ .

*Démonstration.* Par le corolaire 2.15 au théorème de Lagrange, nous avons  $g^{|G|} = e$  et donc en tant qu'opérateur,  $\rho(g)^{|G|} = \mathbb{1}$ . Les valeurs propres de  $\rho(g)$  sont donc des racines de l'unité. Si nous notons  $\lambda_i$  ces valeurs propres, alors  $\chi(g) = \sum_i \lambda_i$ , et en considérant la matrice dans sa base de diagonalisation (lemme de Schur complexe, 12.102), nous voyons que

$$\chi(g^{-1}) = \text{Tr}(\rho(g)^{-1}) = \sum_i \frac{1}{\lambda_i}. \quad (16.60)$$

Mais  $\lambda_i$  étant une racine de l'unité nous avons  $\frac{1}{\lambda_i} = \bar{\lambda}_i$ , ce qui fait que

$$\chi(g^{-1}) = \sum_i \bar{\lambda}_i = \overline{\chi(g)}. \quad (16.61)$$

□

PropJzbfWi

**Proposition 16.18.**

Soient deux représentations irréductibles complexes  $(V, \rho)$  et  $(V', \rho')$  du même groupe fini  $G$ , et  $\chi$  et  $\chi'$  leurs caractères respectifs. Nous avons

(1)  $\langle \chi, \chi' \rangle = 0$  si  $\rho$  et  $\rho'$  ne sont pas équivalentes.

(2)  $\langle \chi, \chi' \rangle = 1$  si les représentations sont équivalentes.

*Démonstration.* Nous considérons les bases  $\{e_1, \dots, e_n\}$  de  $V$  et  $\{f_1, \dots, f_m\}$  de  $V'$ . Puis nous considérons la matrice  $F(k, l) = E_{kl} \in \mathbb{M}_{m,n}(\mathbb{C})$  où pour rappel,  $E_{kl}$  est la matrice de composantes  $(E_{kl})_{ij} = \delta_{ki}\delta_{lj}$ . Nous posons

$$F_G(k, l) = \frac{1}{|G|} \sum_{g \in G} \rho(g) \circ F(k, l) \circ \rho'(g)^{-1}. \quad (16.62)$$

En nous permettant de ne pas réécrire les indices  $k$  et  $l$  de  $F$  et  $F_G$ , nous montrons que  $F_G$  entrelace  $\rho$  et  $\rho'$  :

$$F_G \circ \rho'(t) = \frac{1}{|G|} \sum_{s \in G} \rho(s) \circ F \circ \rho'(s^{-1}) \circ \rho'(t) \quad (16.63a)$$

$$= \frac{1}{|G|} \sum_s \rho(s) F \rho'(s^{-1}t) \quad (16.63b)$$

$$= \frac{1}{|G|} \sum_k \rho(tk) F \rho'(k^{-1}) \quad (16.63c)$$

$$= \frac{1}{|G|} \rho(t) \sum_k \rho(k) F \rho'(k^{-1}) \quad (16.63d)$$

$$= \rho(t) \circ F_G. \quad (16.63e)$$

Dans ce calcul nous avons effectué le changement de variables  $k = (s^{-1}t)^{-1}$  qui donne  $s = tk$ .

Par ailleurs nous avons

$$\left( \rho(g) F(k, l) \rho'(g^{-1}) \right)_{ij} = \sum_{r=1}^n \sum_{s=1}^m \rho(g)_{ir} F(k, l)_{rs} \rho'(g^{-1})_{sj} \quad (16.64a)$$

$$= \sum_{rs} \rho(g)_{ir} \delta_{kr} \delta_{ls} \rho'(g^{-1})_{sj} \quad (16.64b)$$

$$= \rho(g)_{ik} \rho'(g^{-1})_{lj}, \quad (16.64c)$$

et par conséquent

$$F_G(k, l)_{ij} = \frac{1}{|G|} \sum_{g \in G} \rho(g)_{ik} \rho'(g^{-1})_{lj}. \quad (16.65)$$

Si  $\chi$  et  $\chi'$  sont les caractères de  $\rho$  et  $\rho'$ , alors nous avons le produit (16.56) qui donne

$$\langle \chi, \chi' \rangle = \frac{1}{|G|} \sum_{g \in G} \chi(g) \overline{\chi'(g)} \quad (16.66a)$$

$$= \frac{1}{|G|} \sum_g \chi(g) \chi'(g^{-1}) \quad \text{lemme 16.17} \quad (16.66b)$$

$$= \frac{1}{|G|} \sum_g \sum_{i=1}^n \sum_{j=1}^m \rho(g)_{ii} \rho'(g^{-1})_{jj} \quad (16.66c)$$

$$= \sum_{ij} F_G(i, j)_{ij} \quad \text{par (16.65).} \quad (16.66d)$$

Si les représentations  $\rho$  et  $\rho'$  ne sont pas équivalentes, le fait que  $F_G$  en soit un opérateur d'entrelacement implique par le théorème de Schur 16.14 que  $F_G = 0$  et donc  $\langle \chi, \chi' \rangle = 0$ .

Si au contraire les représentations sont équivalentes, alors le lemme 16.16 nous dit que  $\chi = \chi'$  et nous reprenons la définition :

$$\langle \chi, \chi \rangle = \frac{1}{|G|} \sum_g \chi(g) \overline{\chi(g)} = \frac{1}{|G|} \sum_{g \in G} 1 = 1 \quad (16.67)$$

parce que les nombres  $\chi(g)$  sont des racines de l'unité.  $\square$

### 16.2.1 Représentation régulière

DEFooGGPTTooExTxWj

#### Définition 16.19.

Nous notons  $\lambda$  la **représentation régulière gauche**, agissant sur le  $\mathbb{K}$ -espace vectoriel des fonctions  $G \rightarrow \mathbb{K}$  par

$$(\lambda(g)f)(h) = f(g^{-1}h). \quad (16.68)$$

D'autre part nous considérons les fonctions  $\delta_g: G \rightarrow \mathbb{K}$  (ici  $\mathbb{K}$  est  $\mathbb{R}$  ou  $\mathbb{C}$  ou pire) définie par

$$\delta_g(h) = \begin{cases} 1 & \text{si } g = h \\ 0 & \text{sinon.} \end{cases} \quad (16.69)$$

La représentation régulière agit sur les fonctions  $\delta_s$  de la façon suivante :

$$\lambda(g)\delta_s = \delta_{gs} \quad (16.70)$$

parce que  $(\lambda(g)\delta_s)(h) = \delta_s(g^{-1}h) = \delta_{gs}(h)$ .

LEMooGYIKooRRUxHq

#### Lemme 16.20.

Le caractère de la représentation régulière gauche est donné par

$$\chi_\lambda = |G|\delta_e. \quad (16.71)$$

EqUuoVNa

*Démonstration.* Appliquer l'équation (16.71) fonctionne parce que  $\chi_\lambda(e)$  est la dimension de l'espace des fonctions sur  $G$ , c'est-à-dire  $|G|$ . Si par contre  $g \neq e$ , alors  $\lambda(g)$  est une matrice de permutation (dans la base des  $\delta_h$ ) et a donc tous ses éléments diagonaux nuls.  $\square$

Si  $\rho$  est une représentation et si  $f$  est une fonction sur le groupe, alors nous considérons l'opérateur

$$\rho_f = \sum_{g \in G} f(g)\rho(g). \quad (16.72)$$

PropEAXkAY

#### Proposition 16.21 ([554]).

Si  $(\rho, V)$  est une représentation irréductible et si  $f$  est une fonction centrale sur  $G$ , alors l'opérateur  $\rho_f$  est une homothétie de  $V$  de rapport

$$\frac{1}{\dim V} \sum_{g \in G} f(g)\chi(g) \quad (16.73)$$

où  $\chi$  est le caractère de  $\rho$ .

*Démonstration.* Nous commençons par voir que  $\rho_f$  entrelace  $\rho$ . En effet,

$$\rho(t)^{-1} \circ \rho_f \circ \rho(t) = \sum_g f(g)\rho(t^{-1}gt) \quad (16.74a)$$

$$= \sum_h f(tht^{-1})\rho(h) \quad \text{chm.var.h} = t^{-1}gt \quad (16.74b)$$

$$= \sum_h f(h)\rho(h) \quad (16.74c)$$

$$= \rho_f \quad (16.74d)$$

où en écrivant  $f(tht^{-1}) = f(h)$ , nous avons utilisé le fait que  $f$  était centrale. Étant donné que  $\rho_f$  entrelace une représentation irréductible, le lemme de Schur (16.14) nous indique que  $\rho_f$  est une homothétie. Soit  $k$  le facteur d'homothétie. Alors d'une part  $\text{Tr}(\rho_f) = nk$ . D'autre part,

$$\text{Tr}(\rho_f) = \text{Tr} \left( \sum_g f(g)\rho(g) \right) \quad (16.75a)$$

$$= \sum_g f(g) \text{Tr}(\rho(g)) \quad (16.75b)$$

$$= \sum_g f(g)\chi(g). \quad (16.75c)$$



Du coup effectivement

$$k = \frac{1}{n} \sum_{g \in G} f(g) \chi(g). \quad (16.76)$$

□

### 16.2.2 Caractères et représentations : suite et fin

#### Lemme 16.22.

Un groupe fini n'a (à équivalence près) qu'un nombre fini de représentations irréductibles.

*Démonstration.* Les caractères irréductibles forment un système orthonormé (proposition 16.18) et donc libre parmi les fonctions centrales. Donc il y a au plus autant de caractères irréductibles que la dimension de l'espace des fonctions centrales ; et ce dernier est de dimension finie donnée par le nombre de classes de conjugaison de  $G$ . □

Nous savons que les caractères de deux représentations équivalentes sont égaux. Étant donné qu'il n'existe qu'un nombre fini de représentations irréductibles, il existe un nombre fini de caractères irréductibles. Nous pouvons donc fixer les notations suivantes. Les caractères irréductibles seront notés  $\{\varphi_i\}_{i=1,\dots,h}$  et nous noterons  $(\sigma_i, W_i)$  une représentation ayant le caractère  $\varphi_i$ .

#### Théorème 16.23 ([554]).

Soit  $(\rho, V)$  une représentation de  $G$  de caractère  $\chi$ . Alors sa décomposition en représentations irréductibles est donnée par

$$(V, \rho) = \bigoplus_{i=1}^h k_i (W_i, \sigma_i) \quad (16.77)$$

avec  $k_i = \langle \chi, \varphi_i \rangle$ . En particulier, à permutation près des facteurs, la décomposition d'une représentation en représentations irréductibles est unique.

*Démonstration.* La décomposition de  $\chi$  en caractères irréductibles est donnée par  $\chi = \sum_i k_i \varphi_i$  ; en prenant le produit de cette égalité avec  $\varphi_j$  et en tenant compte de l'orthonormalité des caractères irréductibles,

$$\langle \chi, \varphi_j \rangle = \sum_i k_i \langle \varphi_i, \varphi_j \rangle = k_j. \quad (16.78)$$

□

Le théorème suivant est ce qui nous permet de dire que l'étude des caractères et l'étude des représentations, c'est la même chose.

ThoWGkfADd

#### Théorème 16.24.

Soit  $G$  un groupe fini<sup>7</sup>.

ItemZReOWoHi

(1) Deux représentations sont équivalentes si et seulement si elles ont même caractère.

ItemZReOWoHii

(2) Si  $\chi$  est le caractère d'une représentation, alors

(2a)  $\langle \chi, \chi \rangle \in \mathbb{N}$

(2b)  $\langle \chi, \chi \rangle = 1$  si et seulement si la représentation est irréductible.

*Démonstration.* Nous démontrons chaque point séparément.

(1) Le fait que deux représentations équivalentes aient même caractère est le lemme 16.16. Nous montrons l'autre sens. Si  $(\rho, V)$  et  $(\rho', V')$  sont deux représentations irréductibles de décompositions

$$V = \bigoplus_i k_i W_i \quad (16.79a)$$

$$V' = \bigoplus_i k'_i W_i, \quad (16.79b)$$

7. Nous sommes depuis longtemps dans l'étude des représentations des groupes finis.

alors si  $\chi = \chi'$ , nous avons  $k_i = k'_i$  et les représentations sont identiques.

- (2) Soit  $(\rho, V)$  une représentation ayant  $\chi$  comme caractère. En posant  $k_i = \langle \chi, \varphi_i \rangle$  nous avons la décomposition en représentations irréductibles

$$V = \bigoplus_i k_i W_i, \quad (16.80)$$

et aussi

$$\langle \chi, \chi \rangle = \left\langle \sum_i k_i \varphi_i, \sum_j k_j \varphi_j \right\rangle = \sum_i k_i^2 \in \mathbb{N}. \quad (16.81)$$

Ce nombre est de plus égal à 1 si et seulement si tous les termes de la somme sont nuls sauf un qui vaudrait 1. Ce cas donne une représentation irréductible.

□

PropYLnXljk

**Proposition 16.25.**

Si  $(\lambda, R)$  est la représentation régulière gauche de décomposition en représentations irréductibles

$$R = \bigoplus_i k_i W_i, \quad (16.82)$$

alors

(1)  $k_i = \dim W_i$ ,

ITEMooLXIJoDxkGJh

(2)  $\sum_i (\dim W_i)^2 = |G|$ ,

ItemEXAjTih

(3) pour tout  $g \in G$ ,  $\sum_i (\dim W_i) \varphi_i(g) = 0$ <sup>8</sup>.

- (4) Si  $\{(n_i, \varphi_i)\}$  est la liste des couples dimension, caractère des représentations irréductibles non équivalentes, alors pour tout  $s \in G \setminus \{e\}$  nous avons  $\sum_{i=1}^p n_i \varphi_i(s) = 0$  où la somme porte sur les représentations irréductibles non équivalentes.

*Démonstration.* Nous notons  $r$  le caractère de la représentation régulière gauche. Nous avons

$$k_i = \langle r, \varphi_i \rangle = \frac{1}{|G|} \sum_{s \in G} r(s) \overline{\varphi_i(s)} = \overline{\varphi_i(e)}. \quad (16.83)$$

Mais  $\varphi_i(e) = \dim(W_i) \in \mathbb{N}$ , donc nous avons bien  $k_i = \dim(W_i)$ . Le caractère de la représentation régulière peut alors s'exprimer de deux façons :

$$|G| \delta_e = \sum_i (\dim W_i) \varphi_i. \quad (16.84)$$

En évaluant cette égalité en  $e$  nous trouvons directement

$$|G| = \sum_i (\dim W_i)^2, \quad (16.85)$$

et en l'évaluant en  $s \neq e$ , nous trouvons

$$0 = \sum_i (\dim W_i) \varphi_i(s). \quad (16.86)$$

□

fqsdf

Le théorème suivant est valable pour les groupes finis (comme toute cette section).

Thogocemg

**Théorème 16.26** ([554]).

Les caractères irréductibles  $\chi_1, \dots, \chi_h$  forment une base orthonormée des fonctions centrales sur  $G$ .

<sup>8</sup>. Cette propriété est appelée « orthogonalité des colonnes » pour une raison qui apparaîtra au moment de compléter le tableau (16.111).

*Démonstration.* Nous savons déjà qu'ils forment un système orthonormé. Considérons le sous-espace  $H = \text{Span}\{\varphi_i\}_{i=1,\dots,h}$  de l'espace des fonctions centrales sur  $G$ . En vertu de la proposition 4.144, il nous suffit de prouver que  $H^\perp = \{0\}$ . Soit donc  $f$ , une fonction centrale appartenant à  $H^\perp$ . Pour tout  $i$ , nous avons  $\langle f, \varphi_i \rangle = 0$  et donc aussi  $\langle \bar{f}, \bar{\varphi}_i \rangle = 0$ .

Considérant une représentation irréductible  $(\sigma, W)$  de caractère  $\varphi$ , nous savons par la proposition 16.21 que l'opérateur

$$\sigma_{\bar{f}} = \sum_g \bar{f}(g) \varphi(g) \quad (16.87)$$

est une homothétie de rapport  $\langle \bar{f}, \bar{\varphi} \rangle / \dim W = 0$ . Étant donné que toutes les représentations sont des sommes directes de représentations irréductibles, en réalité l'opérateur  $\rho_{\bar{f}}$  est nul pour toute représentation  $\rho$ . En particulier pour la représentation régulière,

$$0 = \lambda_{\bar{f}}(\delta_t) = \sum_{g \in G} \bar{f}(g) \lambda(g)(\delta_t) = \sum_g \bar{f}(g) \delta_{gt}. \quad (16.88)$$

En écrivant cette égalité avec  $t = e$  et puis en appliquant à  $k \in G$  nous trouvons

$$0 = \sum_g \bar{f}(g) \delta_g(k) = \bar{f}(k). \quad (16.89)$$

Donc  $\bar{f} = 0$  et  $f$  est nulle. □

CorbdcVNC

### Corolaire 16.27.

*Le nombre de représentations irréductibles non équivalentes d'un groupe fini est égal à son nombre de classes de conjugaison.*

*Démonstration.* Le nombre de classes de conjugaison est la dimension de l'espace des fonctions centrales qui elle-même est égale au nombre de caractères irréductibles par le théorème 16.26. Enfin deux caractères irréductibles sont égaux si et seulement si les représentations sous-jacentes sont équivalentes. □

CORooWAGXooByreLO

### Corolaire 16.28.

*Toutes les représentations irréductibles d'un groupe abélien sont de dimension 1.*

*Démonstration.* Le corolaire 16.27 nous dit qu'il y a autant de représentations unitaires qu'il n'y a de représentations irréductibles (non équivalentes). Mais les classes de conjugaisons sont des singletons (lemme 1.254). Nous avons donc exactement  $|G|$  représentations irréductibles lorsque  $G$  est abélien.

Mais d'autre part la proposition 16.25(2) donne  $\sum_i (\dim W_i)^2 = |G|$  lorsque la somme parcourt les représentations irréductibles. Il y a  $|G|$  termes à la somme, donc tous les termes doivent être 1. □

## 16.3 Représentation produit tensoriel

### Définition 16.29.

*Soient  $\rho$  et  $\phi$ , deux représentations d'un groupe  $G$  sur des espaces vectoriels  $V$  et  $W$ . La représentation **produit tensoriel** est la représentation*

$$\begin{aligned} \rho \otimes \phi: G &\rightarrow \text{GL}(V \otimes W) \\ (\rho \otimes \phi)(g)(v \otimes w) &= \rho(g)v \otimes \phi(g)w. \end{aligned} \quad (16.90)$$

*L'espace  $V \otimes W$  est le produit tensoriel de  $V$  et  $W$ , défini en 10.193.*

Pour trouver son caractère, nous considérons une base  $\{e_i\}$  de  $V$  et une base  $\{e_\alpha\}$  de  $W$ , et la base  $\{e_i \otimes e_\alpha\}$  de  $V \otimes W$ . Donc

$$(\rho \otimes \phi)(g)(e_i \otimes e_\alpha) = \rho(g)e_i \otimes \phi(g)e_\alpha. \quad (16.91)$$

Nous devons savoir quelle est la composante «  $e_i \otimes e_\alpha$  » de cette dernière expression, et c'est évidemment

$$\rho(g)_{ii} \rho_{\alpha\alpha}, \quad (16.92)$$

ce qui nous amène à dire que

$$\mathrm{Tr}(\rho \otimes \phi)(g) = \sum_i \sum_\alpha \rho(g)_{ii} \phi(g)_{\alpha\alpha} = \mathrm{Tr}(\rho(g)) \mathrm{Tr}(\phi(g)), \quad (16.93)$$

c'est-à-dire au final que

$$\chi_{\rho \otimes \phi} = \chi_\rho \chi_\phi. \quad (16.94)$$

## 16.4 Exemple sur le groupe symétrique

Soit  $G = S_3$ , un des premiers groupes finis non abéliens. On en a une représentation de dimension deux en tant que permutation des sommets d'un triangle équilatéral, donnée dans l'exemple 16.6 ; nous notons  $\rho$  cette représentation.

Nous y avons aussi la représentation de signature donnée par

$$\begin{aligned} \epsilon: S_3 &\rightarrow \mathrm{GL}(\mathbb{C}) \\ \sigma &\mapsto \epsilon(\sigma) \mathrm{Id}. \end{aligned} \quad (16.95)$$

Et enfin il y a la représentation triviale. Ce sont les trois représentations irréductibles ; pour rappel il y a autant de représentations irréductibles que de classes de conjugaison (corolaire 16.27).

| Classe de conjugaison | taille | $\chi_1$ | $\chi_\epsilon$ | $\chi_\rho$ |
|-----------------------|--------|----------|-----------------|-------------|
| Id                    | 1      | 1        | 1               | 2           |
| $(A, B)$              | 3      | 1        | -1              | 0           |
| $(A, B, C)$           | 2      | 1        | 1               | -1          |

Nous calculons par exemple le produit scalaire

$$\langle \chi_1, \chi_\epsilon \rangle = \frac{1}{6} (1 \cdot \chi_1(\mathrm{Id}) \overline{\chi_\epsilon(\mathrm{Id})} + 3 \cdot \chi_1(A, B) \overline{\chi_\epsilon(A, B)} + 2 \cdot \chi_1(A, B, C) \overline{\chi_\epsilon(A, B, C)}) \quad (16.96a)$$

$$= 0. \quad (16.96b)$$

D'autre part nous avons aussi

$$\langle \chi_\rho, \chi_\rho \rangle = \frac{1}{6} (1 \cdot 2 \cdot 2 + 3 \cdot 0 + 2 \cdot 1) = 1. \quad (16.97)$$

## 16.5 Table des caractères du groupe symétrique $S_4$

Pour la table des caractères de  $S_4$ , voir [126]. Et si vous voulez la table des caractères du groupe diédral, vu que ce sont de isométries de  $\mathbb{R}^n$ , il faudra voir plus bas en la section 18.17.

### 16.5.1 Calculs à partir de rien ou presque

Nous savons que les classes de conjugaison dans  $S_4$  sont caractérisées par la structure des décompositions en cycles (proposition 1.278). Elles sont données dans l'exemple 1.281.

Nous avons donc 5 classes de conjugaison, et il nous faut donc 5 représentations irréductibles non équivalentes (corolaire 16.27) dont nous allons chercher les caractères.

La première est la représentation triviale de dimension 1 ; nous notons  $\chi_1$  son caractère et nous avons la ligne

|          |           |    |      |       |        |          |
|----------|-----------|----|------|-------|--------|----------|
| $\chi_1$ | dimension | Id | (12) | (123) | (1234) | (12)(34) |
|          | 1         | 1  | 1    | 1     | 1      | 1        |

(16.98)

Ensuite nous avons la signature qui est un morphisme non trivial  $\epsilon: S_n \rightarrow \{-1, 1\}$ . Nous avons alors la ligne

|                 |           |    |      |       |        |          |
|-----------------|-----------|----|------|-------|--------|----------|
|                 | dimension | Id | (12) | (123) | (1234) | (12)(34) |
| $\chi_\epsilon$ | 1         | 1  | -1   | 1     | -1     | 1        |

EqGNBaxtl  
(16.99)

Une troisième représentation pas trop compliquée à trouver est celle

$$\begin{aligned}\rho_p: S_4 &\rightarrow \text{GL}(4, \mathbb{C}) \\ \rho_p(\sigma)e_i &= e_{\sigma(i)}.\end{aligned}\tag{16.100}$$

Cela n'est pas une représentation irréductible parce que  $\mathbb{C}^4$  se décompose en deux sous-espaces stables :

$$D = \text{Span}(1, 1, 1, 1) \tag{16.101a}$$

$$H = \{x \in \mathbb{C}^4 \text{ tel que } x_1 + x_2 + x_3 + x_4 = 0\}.\tag{16.101b}$$

La représentation induite sur  $D$  est la représentation triviale. Puis sur  $H$ , elle induit une autre représentation que nous allons noter  $\rho_s$ .

Nous allons à présent déduire le caractère de la représentation  $\rho_s$  et prouver qu'elle est irréductible. Il est cependant possible de sauter cette étape en échange d'un certain travail sur les isométries du tétraèdre. Voir la proposition 18.194 et ensuite

Nous avons la décomposition  $\rho_p = \rho_1 \oplus \rho_s$  et donc

$$\chi_p = \chi_1 + \chi_s.\tag{16.102}$$

Nous savons déjà  $\chi_1$ . Le caractère  $\chi_p$  n'est pas très compliqué parce que  $\chi_p(\sigma)$  est une matrice de permutations des vecteurs de base. Donc la matrice  $\rho_p(\sigma)$  a un 1 sur la diagonale pour les  $i$  tels que  $\sigma(i) = i$ . Nous avons donc

$$\chi_p(\text{Id}) = 4 \tag{16.103a}$$

$$\chi_p((12)(34)) = 0 \tag{16.103b}$$

$$\chi_p(1234) = 0.\tag{16.103c}$$

Le caractère  $\chi_s$  peut être calculé par simple soustraction :

|          |           |    |      |       |        |          |
|----------|-----------|----|------|-------|--------|----------|
|          | dimension | Id | (12) | (123) | (1234) | (12)(34) |
| $\chi_s$ | 3         | 3  | 1    | 0     | -1     | -1       |

EqILZsKfo  
(16.104)

Avant d'ajouter cette ligne au tableau des représentations irréductibles nous devons savoir si  $\rho_s$  en est une. Pour cela, tant que nous avons son caractère nous pouvons utiliser le critère du théorème 16.24 :

$$\langle \chi_s, \chi_s \rangle = \frac{1}{|S_4|} \sum_{\sigma \in S_4} \chi_s(\sigma)^2.\tag{16.105}$$

Nous avons tout de suite  $|S_4| = 4 \cdot 3 \cdot 2 = 24$  et puis

$$24\langle \chi_s, \chi_s \rangle = 3^2 + 6 \cdot 1^2 + 8 \cdot 0^2 + 6 \cdot (-1)^2 + 3 \cdot (-1)^2 = 24,\tag{16.106}$$

donc oui, le caractère est irréductible parce que  $\langle \chi_s, \chi_s \rangle = 1$ . Et nous pouvons donc ajouter la ligne (16.104) à notre tableau. Par ailleurs, nous notons qu'elle est de dimension 3.

Pour le reste nous savons qu'il y a autant de représentations irréductibles que de classes de conjugaison, de telle sorte qu'il ne manque que deux représentations irréductibles. De plus la proposition 16.25 nous dit que si  $n_i$  est la dimension de la  $i^{\text{e}}$  représentation irréductible, alors

$$|S_4| = \sum_i n_i^2.\tag{16.107}$$

Dans notre situation, si nous nommons  $n_1$  et  $n_2$  les dimensions des deux représentations qui nous manquent, nous avons  $24 = n_1^2 + n_2^2 + (1^2 + 1^2 + 3^2)$ , c'est-à-dire  $n_1^2 + n_2^2 = 13$ . Il n'y a pas des tonnes de sommes de deux carrés qui font 13. Il y a  $n_1 = 2$  et  $n_2 = 3$ , et c'est tout.

Nous recherchons donc encore une représentation de dimension 2 et une de dimension 3. Pour cela nous allons un peu regarder les produits tensoriels qui s'offrent à nous. Pour faire une dimension 3, il faut faire le produit d'une de dimension 1 par une de dimension 3. Là encore le choix est très limité et nous demande d'essayer

$$\rho_W = \rho_s \otimes \rho_\epsilon \quad (16.108)$$

qui agit sur l'espace  $V_2 \otimes V_\epsilon$  par

$$\rho_W(g)(v \otimes x) = \rho_s(g)v \otimes \rho_\epsilon(g)x. \quad (16.109)$$

Pour savoir son caractère nous utilisons la petite formule toute simple (16.94) : nous multiplions case par case les tableaux (16.104) et (16.99) :

$$\begin{array}{c|c|c|c|c|c|c} & \text{dimension} & \text{Id} & (12) & (123) & (1234) & (12)(34) \\ \hline \chi_W & 3 & 3 & -1 & 0 & 1 & -1 \end{array} \quad (16.110)$$

Avant de réellement ajouter cette ligne au tableau, nous devons nous assurer qu'elle est bien irréductible. Nous utilisons le même critère :  $\langle \chi_W, \chi_W \rangle = 1$ , donc c'est bon.

Pour trouver le dernier caractère, que nous nommerons  $\chi_u$ , il ne faut pas beaucoup d'imagination. Il suffit d'utiliser les relations d'orthogonalité du théorème 16.26, en sachant que la dimension est 2 et qu'alors  $\chi_u(\text{Id}) = 2$ , c'est pas trop compliqué :

$$\begin{array}{c|c|c|c|c|c|c} & \text{dimension} & \text{Id} & (12) & (123) & (1234) & (12)(34) \\ \hline \chi_1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline \chi_\epsilon & 1 & 1 & -1 & 1 & -1 & 1 \\ \hline \chi_s & 3 & 3 & 1 & 0 & -1 & -1 \\ \hline \chi_W & 3 & 3 & -1 & 0 & 1 & -1 \\ \hline \chi_u & 2 & 2 & b & c & d & e \end{array} \quad \begin{array}{l} \text{EqOKtZYEQ} \\ (16.111) \end{array}$$

Les relations d'orthogonalité des colonnes de la propriété 16.25 nous permettent de calculer les coefficients manquants. En pratique, il suffit de prendre le produit scalaire de chaque ligne avec la première et d'égaliser avec zéro. Nous trouvons  $b = 0$ ,  $c = 1$ ,  $d = 0$ , et  $e = 2$ . Le tableau final est :

$$\begin{array}{c|c|c|c|c|c|c} & \text{dimension} & \text{Id} & (12) & (123) & (1234) & (12)(34) \\ \hline \chi_1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline \chi_\epsilon & 1 & 1 & -1 & 1 & -1 & 1 \\ \hline \chi_s & 3 & 3 & 1 & 0 & -1 & -1 \\ \hline \chi_W & 3 & 3 & -1 & 0 & 1 & -1 \\ \hline \chi_u & 2 & 2 & 0 & -1 & 0 & 2 \end{array} \quad (16.112)$$

Notons que nous sommes parvenus à remplir la dernière ligne sans rien savoir de la représentation qui va avec. Nous allons cependant donner une interprétation géométrique et fixer cette représentation comme agissant sur le triangle équilatéral en 16.33.

### 16.5.2 À propos de la représentation $\rho_s$

Une des représentations trouvées (la représentation  $\rho_s$ ) peut être vue comme le groupe  $\text{Iso}(T)$  des isométries affine du tétraèdre<sup>9</sup> grâce à la proposition 18.194 qui donne un isomorphisme de groupe  $S_4 \simeq \text{Iso}(T)$  lorsque  $T$  est un tétraèdre régulier de  $\mathbb{R}^3$ .

Nous verrons donc ça plus en détail dans la section 18.10.18.

9. Définition 12.180.

### 16.5.3 À propos de la représentation $\rho_u$

Nous nous penchons à présent sur la représentations  $\rho_u$  dont nous ne savons rien à part qu'elle est de dimension 2 et son caractère.

#### Lemme 16.30.

Nous avons  $\rho_u(s) = \text{Id}$  pour tout  $s \in V_4$ .

*Démonstration.* Tous les éléments de  $V_4$  sont conjugués (à part l'identité, mais pour elle le résultat est clair), donc il suffit de prouver le résultat pour un élément quelconque.

L'endomorphisme  $\rho_u((12)(34))$  est un endomorphisme d'ordre 2 sur  $\mathbb{R}^2$  dont la trace est 2. Imposons donc

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (16.113)$$

sous la contrainte  $a + d = 2$ . La résolution est assez rapide et donne  $b = c = 0$ ,  $a = d = 1$ .

Vous voulez une démonstration plus technologique ? Oui ? Alors commencez par remarquer que l'opérateur  $A = \rho_u((12)(34))$  vérifie  $A^2 = 1$ , donc le polynôme  $X^2 - 1$  est un polynôme annulateur de  $A$ . Il est peut-être minimal ou peut être pas, mais en tout cas le polynôme minimal divise celui-là et donc est soit  $X - 1$  soit  $X + 1$  soit  $X^2 - 1$ . Dans les trois cas il est scindé à racines simples, et l'endomorphisme  $A$  est diagonalisable par le théorème 8.218(3).

Mais comme  $A^2 = 1$ , les valeurs propres (ce qui est sur la diagonale) de  $A$  ne peuvent être que  $\pm 1$ . La trace étant 2, les éléments diagonaux ne peuvent être que 1. Et  $A = \text{Id}$ .  $\square$

Le groupe  $V_4$  définit en 5.56 est normal dans  $S_4$ , donc le quotient  $S_4/V_4$  est un groupe par le lemme 2.11.

#### Lemme 16.31.

L'application

$$\begin{aligned} \tilde{\rho}_u: S_4/V_4 &\rightarrow \text{GL}(2, \mathbb{R}) \\ [g] &\mapsto \rho_u(g) \end{aligned} \quad (16.114)$$

est bien définie et donne une représentation irréductible de  $S_4/V_4$ .

*Démonstration.* Montrons que c'est bien défini. Si  $s \in V_4$  nous devons prouver que  $\rho_u(gs) = \rho_u(g)$ . Vu que  $\rho_u$  est un homomorphisme (c'est une représentation), et que  $\rho_u(s) = \text{Id}$  nous avons directement

$$\rho_u(gs) = \rho_u(g)\rho_u(s) = \rho_u(g). \quad (16.115)$$

Nous devons prouver que la représentation  $\tilde{\rho}_u$  est irréductible. Si un sous-espace non trivial  $\text{Span}(x)$  était stabilisé par  $\tilde{\rho}_u$ , il serait également stabilisé par  $\rho_u$ . Mais comme  $\rho_u$  est irréductible, elle ne stabilise personne.  $\square$

#### Lemme 16.32.

Le groupe  $S_4/V_4$  est un groupe non-abélien, isomorphe à  $S_3$ .

*Démonstration.* Le groupe  $S_4/V_4$  a une représentation irréductible de dimension 2, et n'est donc pas abélien par le corolaire 16.28.

Il contient  $|S_4|/|V_4| = 24/4 = 6$  éléments (théorème de Lagrange 2.14). Or  $6 = 3 \times 2$ , donc le groupe  $S_4/V_4$  est dans le cas non-abélien du théorème 5.25(2). Cette partie parle d'unicité du groupe non-abélien d'ordre 6. Or  $S_3$  est un groupe non-abélien d'ordre 6, donc  $S_4/V_4$  est isomorphe à  $S_3$ .  $\square$

Attention : il n'est pas correct de dire que  $S_4/V_4$  est un sous-groupe de  $S_4$  juste parce que c'est un quotient de  $S_4$  ; ce n'est en général pas vrai (exemple 2.12).

NORMooQQCYooILyOxc

#### 16.33.

Nous sommes maintenant aptes à identifier la représentation  $\rho_u$ . D'abord nous nous rappelons de la

représentation  $\rho_s: S_4 \rightarrow \text{Iso}(T)$  de  $S_4$  sur le tétraèdre. Ensuite si  $A$  est un sommet dudit tétraèdre et que  $S_3 \subset S_4$  est la partie qui fixe  $A$  alors nous avons une représentation

$$\rho_s: S_3 \rightarrow \text{Iso}(T) \quad (16.116)$$

qui agit en réalité sur le triangle équilatéral  $T'$  opposé au sommet  $A$ .

Nous avons finalement la chaîne d'homomorphismes de groupes

$$S_4 \xrightarrow{\text{proj}} S_4/V_4 \xrightarrow{\cong} S_3 \xrightarrow{\rho_s} \text{Iso}(T') \quad (16.117)$$

Cela est donc une représentation  $S_4 \rightarrow \text{Iso}(T')$ . Elle est de dimension 2 et est irréductible (elle contient les rotations d'angle  $2\pi/3$  qui ne fixent aucune direction). Elle est donc la représentation  $\rho_u$  qui est la seule irréductible de dimension 2.

Nous avons donc montré que la représentation  $\rho_u$  dont nous ne savions rien est la représentation de  $S_3$  sur un triangle équilatéral obtenue à partir de celle de  $S_4$  sur le tétraèdre, en fixant un point.





# Chapitre 17

## Encore de l'analyse (et c'est pas fini)

### 17.1 Densité des polynômes

CorRSzQD

#### Corolaire 17.1.

Si  $X \subset \mathbb{R}$  est compact et de mesure finie<sup>1</sup>, alors l'ensemble des polynômes est dense dans  $(C(X, \mathbb{R}), \|\cdot\|_2)$ .

*Démonstration.* Si  $f$  est une fonction dans  $C(X, \mathbb{R})$  et si  $\epsilon \geq 0$  est donné alors nous pouvons considérer un polynôme  $P$  tel que  $\|f - P\|_\infty \leq \epsilon$ . Dans ce cas nous avons

$$\|f - P\|_2^2 = \int_X |f(x) - P(x)|^2 dx \leq \int_X \epsilon^2 dx = \epsilon^2 \mu(X) \quad (17.1)$$

où  $\mu(X)$  est la mesure de  $X$  (finie par hypothèse). □

### 17.2 Primitive et intégrale

Nous avons déjà parlé de primitive de fonction continue en la proposition 12.450.

PropHFWNpRb

#### Proposition 17.2.

Soit  $I$  un intervalle borné ouvert de  $\mathbb{R}$ . Une fonction  $h \in C_c^\infty(I)$  admet une primitive dans  $C_c^\infty(I)$  si et seulement si  $\int_I h = 0$ .

*Démonstration.* Si une primitive  $H$  de  $h$  est à support compact, alors

$$\int_I h = H(b) - H(a) = 0 - 0 = 0. \quad (17.2)$$

Pas de problèmes dans ce sens.

Supposons maintenant que  $\int_I h = 0$ . Le fait que  $h$  admette une primitive dans  $C^\infty(I)$  est évident : toute fonction continue admet une primitive<sup>2</sup>. Soit  $H$  une telle primitive et  $\tilde{H} = H - H(b)$ . Alors  $\tilde{H}(b) = 0$  et

$$\tilde{H}(a) = H(a) - H(b) = - \int_I h = 0. \quad (17.3)$$

Nous rappelons que le support d'une fonction est la fermeture de l'ensemble des points de non-annulation.

Supposons que le support de  $h$  soit inclus dans  $[m, M] \subset ]a, b[$ . En prenant des nombres  $m'$  et  $M'$  tels que  $a < m' < m$  et  $M < M' < b$  (nous insistons sur le caractère strict de ces inégalités), la fonction  $h$  est nulle sur  $[a, m']$  et sur  $[M', b]$  ; la fonction  $\tilde{H}$  doit donc y être constante. Mais nous avons déjà vu que  $\tilde{H}(a) = \tilde{H}(b) = 0$ . Donc l'ensemble des points sur lesquels  $\tilde{H}$  n'est pas nul est inclus dans  $]m', M'[$  et donc est strictement (des deux côtés) inclus dans  $I$ . □

1. Dans  $\mathbb{R}$  cette hypothèse est évidemment superflue par rapport à l'hypothèse de compacité ; mais ça suggère des généralisations ...

2. Théorème 12.450.

### 17.2.1 Théorème taubérien de Hardy-Littlewood

Un théorème **taubérien** est un théorème qui compare les modes de convergence d'une série.

#### Lemme 17.3.

Si  $f$  et  $g$  sont des fonctions continues, alors  $s(x) = \max\{f(x), g(x)\}$  est également une fonction continue.

*Démonstration.* Soit  $x_0$  et prouvons que  $s$  est continue en  $x_0$ . Si  $f(x_0) \neq g(x_0)$  (supposons  $f(x_0) > g(x_0)$  pour fixer les idées), alors nous avons un voisinage de  $x_0$  sur lequel  $f > g$  et alors  $s = f$  sur ce voisinage et la continuité provient de celle de  $f$ .

Si au contraire  $f(x_0) = g(x_0) = s(x_0)$  alors si  $(a_n)$  est une suite tendant vers  $x_0$ , nous prenons  $N$  tel que  $|f(a_n) - f(x_0)| \leq \epsilon$  pour tout  $n > N$  et  $M$  tel que  $|g(a_n) - g(x_0)| \leq \epsilon$  pour tout  $n > M$ . Alors pour tout  $n > \max\{N, M\}$  nous avons

$$|s(a_n) - s(x_0)| \leq \epsilon, \quad (17.4)$$

d'où la continuité de  $s$  en  $x_0$ .  $\square$

La proposition suivante dit que si une fonction connaît un saut, alors on peut la lisser par une fonction continue.

PropTleYVw

#### Proposition 17.4.

Soit  $f$  continue sur  $[a, x_0[$  et sur  $[x_0, b]$  avec  $f(x_0^-) < f(x_0)$ . En particulier nous supposons que  $f(x_0^-)$  existe et est finie. Alors pour tout  $\epsilon > 0$ , il existe une fonction continue  $s$  telle que sur  $[a, b]$  on ait  $s \geq f$  et

$$\int_a^b (s - f)(x) dx \leq \epsilon. \quad (17.5)$$

*Démonstration.* Nous notons  $A$  la taille du saut :

$$A = f(x_0) - f(x_0^-). \quad (17.6)$$

Quitte à changer  $a$  et  $b$ , nous pouvons supposer que

$$f(x) < f(x_0) + \frac{A}{3} \quad (17.7)$$

pour  $x \in [a, x_0[$  et

$$f > f(x_0) + \frac{2A}{3} \quad (17.8)$$

pour  $x \in [x_0, b]$ . C'est le théorème des valeurs intermédiaires qui nous permet de faire ce choix.

Soit  $m(x)$  la droite qui joint le point  $(x_0 - \epsilon, f(x_0 - \epsilon))$  au point  $(x_0, f(x_0^+))$ . Nous posons

$$s(x) = \begin{cases} f(x) & \text{si } x < x_0 - \epsilon \\ \max\{m(x), f(x)\} & \text{si } x_0 - \epsilon \leq x \leq x_0 \\ f(x) & \text{si } x > x_0. \end{cases} \quad (17.9)$$

En vertu des différents choix effectués, c'est une fonction continue. En effet

$$s(x_0 - \epsilon) = \max\{f(x_0 - \epsilon), f(x_0, \epsilon)\} = f(x_0 - \epsilon) \quad (17.10)$$

et

$$s(x_0) = \max\{m(x_0), f(x_0^+)\} = f(x_0^+) \quad (17.11)$$

parce que  $m(x_0) = f(x_0^+)$ . En ce qui concerne l'intégrale, si nous posons

$$M = \sup_{x, y \in [a, b]} |f(x) - f(y)|, \quad (17.12)$$

nous avons

$$\int_a^b s - f = \int_{x_0 - \epsilon}^{x_0} s - f \leq \epsilon M. \quad (17.13)$$

$\square$

**Lemme 17.5.**

Pour tout polynôme  $P$ , nous avons la formule

$$\lim_{x \rightarrow 1^-} (1-x) \sum_{n=0}^{\infty} x^n P(x^n) = \int_0^1 P(x) dx. \quad (17.14)$$

*Démonstration.* D'abord pour  $P = 1$ , la formule se réduit à la série géométrique connue (proposition 10.159). Ensuite nous prouvons la formule pour le polynôme  $P = X^k$  et la linéarité fera le reste pour les autres polynômes. Nous avons

$$(1-x) \sum_n x^n x^{kn} = (1-x) \sum_n (x^{1+k})^n = \frac{1-x}{1-x^{1+k}} = \frac{1}{1+x+\dots+x^k}. \quad (17.15)$$

Donc

$$\lim_{x \rightarrow 1^-} (1-x) \sum_n x^n P(x^n) = \frac{1}{1+k}. \quad (17.16)$$

Par ailleurs, c'est vite vu que

$$\int_0^1 x^k dx = \frac{1}{k+1}. \quad (17.17)$$

□

**Théorème 17.6** (Hardy-Littlewood[555]).

Soit  $(a_n)$  une suite réelle telle que

- (1)  $\frac{a_n}{n}$  tend vers une constante,
- (2)  $F(x) = \sum_{n=0}^{\infty} a_n x^n$  a un rayon de convergence  $\geq 1$ ,
- (3)  $\lim_{x \rightarrow 1^-} F(x) = l$ .

Alors  $\sum_{n=0}^{\infty} a_n = l$ .

*Démonstration.* Quitte à prendre la suite  $b_0 = a_0 - l$  et  $b_n = a_n$ , on peut supposer  $l = 0$ .

Soit  $\Gamma$  l'ensemble des fonctions

$$\gamma: [0, 1] \rightarrow \mathbb{R} \quad (17.18)$$

telles que

- (1)  $\sum_{n=0}^{\infty} a_n \gamma(x^n)$  converge pour  $0 \leq x < 1$ ,
- (2)  $\lim_{x \rightarrow 1^-} \sum_{n \geq 0} a_n \gamma(x^n) = 0$ .

Ce  $\Gamma$  est un espace vectoriel.

- (1) Les polynômes sont dans  $\Gamma$  Soit  $\gamma(t) = t^s$ . Pour  $0 \leq x < 1$  nous avons

$$\sum_{n=0}^{\infty} a_n \gamma(x^n) = \sum_{n=0}^{\infty} a_n x^{ns} < \sum_{n=0}^{\infty} a_n x^n. \quad (17.19)$$

Donc la condition de convergence est vérifiée. En ce qui concerne la limite,

$$\lim_{x \rightarrow 1^-} \sum_{n=0}^{\infty} a_n x^{ns} = \lim_{x \rightarrow 1^-} F(x^s) = 0 \quad (17.20)$$

parce que par hypothèse,  $\lim_{x \rightarrow 1^-} F(x) = 0$ .

(2) **Définition de la fonction qui va donner la réponse** Nous considérons la fonction

$$g(t) = \begin{cases} 0 & \text{si } 0 \leq t < 1/2 \\ 1 & \text{si } 1/2 \leq t \leq 1, \end{cases} \quad (17.21)$$

c'est-à-dire  $g = \mathbb{1}_{[1/2, 1]}$ . Nous montrons que si  $g \in \Gamma$ , alors le théorème est terminé. Si  $0 \leq x \leq 1$ , on a  $0 \leq x^n < 1/2$  dès que

$$n > -\frac{\ln(2)}{\ln(x)} \quad (17.22)$$

avec une note comme quoi  $\ln(x) < 0$ , donc la fraction est positive. Nous désignons par  $N_x$  la partie entière de ce  $n$  adapté à  $x$ . L'idée est que la fonction  $g(x^n)$  est la fonction indicatrice de  $0 \leq n \leq N_x$ , et donc

$$\sum_{n \geq 0} a_n g(x^n) = \sum_{n=0}^{N_x} a_n. \quad (17.23)$$

Mais si  $x \rightarrow 1^-$ , alors  $N_x \rightarrow \infty$ , donc

$$\lim_{N \rightarrow \infty} \sum_{n=0}^N a_n = \lim_{x \rightarrow 1^-} \sum_{n=0}^{N_x} a_n = \lim_{x \rightarrow 1^-} \sum_{n \in \mathbb{N}} a_n g(x^n), \quad (17.24)$$

et cela fait zéro si  $g \in \Gamma$ .

(3) **Approximation de  $g$  par des polynômes** Nous considérons la fonction

$$h(t) = \frac{g(t) - t}{t(1-t)} = \begin{cases} \frac{1}{t-1} & \text{si } t \in [0, 1/2[ \\ \frac{1}{t} & \text{si } t \in [1/2, 1]. \end{cases} \quad (17.25)$$

La seconde égalité est au sens du prolongement par continuité. La fonction  $h$  est une fonction non continue qui fait un saut de  $-2$  à  $2$  en  $x = 1/2$ . En vertu de la proposition 17.4 (un peu adaptée), nous pouvons considérer deux fonctions continues  $s_1$  et  $s_2$  telles que

$$s_1 \leq h \leq s_2 \quad (17.26)$$

et

$$\int_0^1 s_2 - s_1 \leq \epsilon. \quad (17.27)$$

Notons que l'inégalité  $s_1 \leq s_2$  doit être stricte sur au moins un petit intervalle autour de  $x = 1/2$ . Soient  $P_1$  et  $P_2$ , deux polynômes tels que  $\|P_1 - s_1\|_\infty \leq \epsilon$  et  $\|P_2 - s_2\|_\infty \leq \epsilon$  (ici la norme supremum est prise sur  $[0, 1]$ ). C'est le théorème de Stone-Weierstrass (12.446) qui nous permet de le faire.

Nous posons aussi<sup>3</sup>

$$Q_1 = P_1 + \epsilon \quad (17.28a)$$

$$Q_2 = P_2 - \epsilon. \quad (17.28b)$$

Nous avons

$$\int_0^1 Q_1 - Q_2 \leq \int_0^1 Q_1 - P_1 + P_1 - P_2 + P_2 - Q_2. \quad (17.29)$$

Pour majorer cela, d'abord  $Q_1 - P_1 = P_2 - Q_2 = \epsilon$ , ensuite,

$$P_1 - P_2 = P_1 - s_1 + s_1 - s_2 + s_2 - P_2 \quad (17.30)$$

dans lequel nous avons  $P_1 - s_1 \leq \epsilon$ ,  $s_2 - P_2 \leq \epsilon$  et  $\int_0^1 s_1 - s_2 \leq \epsilon$ . Au final, nous posons  $q = Q_2 - Q_1$  et nous avons

$$\int_0^1 q \leq 5\epsilon. \quad (17.31)$$

---

3. À ce niveau, je crois qu'il y a une faute de frappe dans [555].

Enfin nous posons aussi

$$R_i(x) = x + x(1-x)Q_i. \quad (17.32)$$

Ces polynômes vérifient  $R_i(0) = 0$ ,  $R_i(1) = 1$  et

$$R_1 \leq g \leq R_2 \quad (17.33)$$

parce que

$$Q_1 \leq P_1 \leq h \leq P_2 \leq Q_2 \quad (17.34)$$

et

$$t + t(1-t)Q_1 \leq \underbrace{t + t(1-t)h(t)}_{g(t)} \leq t + t(1-t)Q_2. \quad (17.35)$$

- (4) **Preuve que  $g$  est dans  $\Gamma$**  D'abord si  $0 \leq x < 1$ ,  $x^N < \frac{1}{2}$  pour un certain  $N$ , et alors  $g(x^N) = 0$ . Du coup la série

$$\sum_{n=0}^{\infty} a_n g(x^n) = \sum_{n=0}^N a_n \quad (17.36)$$

est une somme finie qui converge donc.

D'autre part nous prenons  $M$  tel que  $|a_n| < \frac{M}{n}$  pour tout  $n$ . Nous majorons  $\sum_{n \in \mathbb{N}} a_n g(x^n)$  en utilisant  $R_1$ . Mais vu que  $R_1$  est un polynôme, nous pouvons dire que  $|\sum_{n=0}^{\infty} a_n R_1(x^n)| \leq \epsilon$  en prenant  $x \in [\lambda, 1[$  et  $\lambda$  assez grand. Nous avons :

$$\left| \sum_{n=0}^{\infty} a_n g(x^n) \right| \leq \left| \sum_{n=0}^{\infty} a_n g(x^n) - \sum_{n=0}^{\infty} a_n R_1(x^n) \right| + \underbrace{\left| \sum_{n=0}^{\infty} a_n R_1(x^n) \right|}_{\leq \epsilon} \quad (17.37a)$$

$$\leq \epsilon + \sum_{n=0}^{\infty} |a_n| (g - R_1)(x^n) \quad (17.37b)$$

$$\leq \epsilon + \sum_{n=0}^{\infty} |a_n| (R_2 - R_1)(x^n) \quad (17.37c)$$

$$\leq \epsilon + M \sum_{n=0}^{\infty} \frac{x^n(1-x^n)}{n} (Q_2 - Q_1)(x^n) \quad (17.37d)$$

$$= \epsilon + M \sum_{n=0}^{\infty} \frac{x^n(1-x^n)}{n} q(x^n) \quad (17.37e)$$

$$\leq \epsilon + M(1-x) \sum_n x^n q(x^n). \quad (17.37f)$$

Justifications :

- La ligne (17.37d) vient du fait que  $R_2 - R_1 = x(1-x)(Q_2 - Q_1)$ .
- La ligne (17.37f) provient d'une majoration sauvage de  $1/n$  par 1 et de  $1 - x^n$  par  $1 - x$ .

Par le lemme 17.5, nous avons alors

$$\lim_{x \rightarrow 1^-} \left| \sum_n a_n g(x^n) \right| \leq \epsilon + M \int_0^1 q \leq 6\epsilon. \quad (17.38)$$

□

### 17.2.2 Théorème de Müntz

ThoAEYDdHp

**Théorème 17.7** (Théorème de Müntz[556, 557, 558]).

Soit  $C_0([0, 1])$ , l'espace des fonctions continues sur  $[0, 1]$  muni de la norme  $\|\cdot\|_\infty$  ou  $\|\cdot\|_2$  et une suite  $(\alpha_n)$  strictement croissante de nombres positifs. Nous notons  $\phi_\lambda$  la fonction  $x \mapsto x^\lambda$ .

Alors

$$\overline{\text{Span}\{1, \phi_{\alpha_n}\}} \quad (17.39)$$

est dense dans  $C_0([0, 1])$  si et seulement si

$$\sum_{n=2}^{\infty} \frac{1}{\alpha_n} = +\infty. \quad (17.40)$$

Nous prouvons le théorème pour la norme  $\|\cdot\|_2$ .

*Démonstration.* Soit  $m \in \mathbb{R}^+$ ; nous notons  $\Delta_N(m)$  la distance entre  $\phi_m$  et  $\text{Span}\{\phi_{\alpha_1}, \dots, \phi_{\alpha_N}\}$ . Cette distance peut être évaluée avec le déterminant de Gram (proposition 8.20)

$$\Delta_N(m)^2 = \frac{G(\phi_m, \phi_{\alpha_1}, \dots, \phi_{\alpha_N})}{G(\phi_{\alpha_1}, \dots, \phi_{\alpha_N})}. \quad (17.41)$$

Pour calculer cela, nous avons besoin des produits scalaires<sup>4</sup>

$$\langle \phi_a, \phi_b \rangle = \int_0^1 x^{a+b} dx = \frac{1}{a+b+1}. \quad (17.42)$$

Pour avoir des notations plus compactes, nous notons  $\alpha_0 = m$ . Donc nous avons à calculer le déterminant

$$G(\phi_m, \phi_{\alpha_1}, \dots, \phi_{\alpha_N}) = \det \left( \frac{1}{\alpha_i + \alpha_j + 1} \right) \quad (17.43)$$

où  $i, j = 0, \dots, N$ . Nous reconnaissons un déterminant de Cauchy (proposition 8.21) en posant, dans  $\frac{1}{\alpha_i + \alpha_j + 1}$ ,  $a_i = \alpha_i$  et  $b_j = \alpha_j + 1$ . Étant donné que  $b_j - b_i = a_j - a_i$ , nous avons

$$G(\phi_m, \phi_{\alpha_1}, \dots, \phi_{\alpha_N}) = \frac{\prod_{0 \leq i < j \leq N} (\alpha_j - \alpha_i)^2}{\prod_{i=0}^N \prod_{j=0}^N (\alpha_i + \alpha_j + 1)}. \quad (17.44)$$

Nous séparons maintenant les termes où  $i$  ou  $j$  sont nuls. En ce qui concerne le dénominateur, il faut prendre tous les couples  $(i, j)$  avec  $i$  et  $j$  éventuellement égaux à zéro. Nous décomposant cela en trois paquets. Le premier est  $(0, 0)$ ; le second est  $(0, i)$  (chaque couple arrive en fait deux fois parce qu'il y a aussi  $(i, 0)$ ); et le troisième sont les  $i, j$  tous deux différents de zéro :

$$(2m+1) \prod_{ij} (\alpha_i + \alpha_j + 1) \prod_i (\alpha_i + m + 1)^2. \quad (17.45)$$

Notons que dans le produit central, le carré est contenu dans le fait qu'on écrit  $\prod_{ij}$  et non  $\prod_{i < j}$ . Nous avons donc

$$G(\phi_m, \phi_{\alpha_1}, \dots, \phi_{\alpha_N}) = \frac{\prod_{i < j} (\alpha_i - \alpha_j)^2 \prod_i (\alpha_i - m)^2}{(2m+1) \prod_{ij} (\alpha_i + \alpha_j + 1) \prod_i (\alpha_i + m + 1)^2}. \quad (17.46)$$

Le calcul de  $G(\phi_{\alpha_1}, \dots, \phi_{\alpha_N})$  est plus simple<sup>5</sup> :

$$G(\phi_{\alpha_1}, \dots, \phi_{\alpha_N}) = \frac{\prod_{i < j} (\alpha_i - \alpha_j)^2}{\prod_{ij} (\alpha_i + \alpha_j + 1)}. \quad (17.47)$$

4. C'est ici qu'on se particularise à la norme  $\|\cdot\|_2$ .

5. Je crois qu'il y a une faute de frappe dans le dénominateur de [556].

En divisant l'un par l'autre il ne reste que les facteurs comprenant  $m$  et en prenant la racine carrée,

$$\Delta_N(m) = \frac{1}{\sqrt{2m+1}} \prod_{i=1}^N \left| \frac{\alpha_i - m}{\alpha_i + m + 1} \right|. \quad (17.48) \quad \text{EqANiuNB}$$

Nous passons maintenant à la preuve proprement dite. Supposons que  $V = \text{Span}\{\phi_{\alpha_i}, i \in \mathbb{N}\}$  est dense. Si  $m$  est un des  $\alpha_i$ , il peut évidemment être approché par les  $\phi_{\alpha_i}$ . Mais vue la densité de  $V$ , un  $\phi_m$  avec  $m \neq \alpha_i$  (pour tout  $i$ ) alors  $\phi_m$  peut également être arbitrairement approché par les  $\phi_{\alpha_i}$ , c'est-à-dire que

$$\lim_{N \rightarrow \infty} \Delta_N(m) = 0. \quad (17.49)$$

Nous posons

$$u_n = \ln \left( \frac{\alpha_n - m}{\alpha_n + m + 1} \right) \quad (17.50)$$

et nous prouvons que la série  $\sum_n u_n$  diverge. En effet nous nous souvenons de la formule  $\ln(ab) = \ln(a) + \ln(b)$ , de telle sorte que la  $N^{\text{e}}$  somme partielle de  $\sum_n u_n$  est

$$\ln \left( \frac{\alpha_1 - m}{\alpha_1 + m + 1} \cdot \dots \cdot \frac{\alpha_N - m}{\alpha_N + m + 1} \right) = \ln(\sqrt{2m+1} \Delta_N(m)), \quad (17.51)$$

qui tend vers  $-\infty$  lorsque  $N \rightarrow \infty$ .

Si la suite  $(\alpha_n)$  est majorée et plus généralement si nous n'avons pas  $\alpha_n \rightarrow \infty$ , alors évidemment la série  $\sum_n \frac{1}{\alpha_n}$  diverge. Nous supposons donc que  $\lim_{n \rightarrow \infty} \alpha_n = \infty$ . Nous avons aussi <sup>6</sup>

$$u_n = \ln \left( \frac{\alpha_n - m}{\alpha_n + m + 1} \right) = \ln \left( 1 - \frac{2m+1}{\alpha_n + m + 1} \right) \sim -\frac{2m+1}{\alpha_n}. \quad (17.52)$$

Une justification est donnée à l'équation (15.373). Ce que nous avons surtout est

$$\sum_n u_n \sim -(2m+1) \sum_n \frac{1}{\alpha_n}. \quad (17.53)$$

Étant donné que la série de gauche diverge, celle de droite diverge <sup>7</sup>.

Nous prouvons maintenant le sens opposé : nous supposons que la série  $\sum_n 1/\alpha_n$  diverge et nous posons

$$V = \text{Span}\{\phi_{\alpha_n} \text{ tel que } n \in \mathbb{N}\}. \quad (17.54)$$

Il suffit de prouver que  $\phi_m \in \bar{V}$  pour tout  $m$  parce qu'un corolaire du théorème de Stone-Weierstrass 17.1 montre que  $\text{Span}\{\phi_k \text{ tel que } k \in \mathbb{N}\}$  est dense dans  $C$  pour la norme  $\|\cdot\|_2$ .

Si  $\alpha_n \rightarrow \infty$ , nous avons :

$$u_n \sim \frac{2m+1}{\alpha_n} \rightarrow 0 \quad (17.55)$$

et alors  $\Delta_N(m) \rightarrow 0$ . Dans ce cas nous avons immédiatement  $\phi_m \in \bar{V}$ .

Si par contre  $\alpha_n$  ne tend pas vers l'infini, nous repartons de l'expression (17.48), nous posons  $0 < \alpha = \sup_i \alpha_i$  et nous calculons :

$$\sqrt{2m+1} \Delta_N(m) = \prod_{i=1}^N \frac{|\alpha_i - m|}{\alpha_i + m + 1} \quad (17.56a)$$

$$\leq \prod_{i=1}^N \frac{\alpha_i + m}{\alpha_i + m + 1} = \prod_{i=1}^N \left( 1 - \frac{1}{\alpha_i + m + 1} \right) \quad (17.56b)$$

$$\leq \prod_{i=1}^N \left( 1 - \frac{1}{\alpha + m + 1} \right) = \left( 1 - \frac{1}{\alpha + m + 1} \right)^N. \quad (17.56c)$$

Cette dernière expression tend vers 0 lorsque  $N \rightarrow \infty$ . □

6. Je crois qu'il y a une faute de signe dans la dernière expression de [557].

7. Nous utilisons le fait que si  $u_n = \sum v_n$  en tant que suites et si  $\sum_n u_n$  diverge, alors  $\sum_n v_n$  diverge.



REMooGPYYooCQJwFa

**Remarque 17.8.**

Certaines sources<sup>8</sup> citent le théorème de Müntz comme ceci (avec un implicite que  $\alpha_i \neq 0$ ) :

$$\overline{\text{Span}\{1, \phi_{\alpha_i}\}} = C([0, 1]) \Leftrightarrow \sum_{i \geq 1} \frac{1}{\alpha_i} = +\infty. \quad (17.57)$$

EQooPCSZooUDSzwQ

Que penser de la présence explicite du 1 (c'est-à-dire de  $\phi_0$ ) ou non dans l'ensemble ?

Première chose : la présence éventuelle de  $\phi_0$  est la raison pour laquelle nous faisons commencer la somme à  $i = 2$  et non  $i = 1$ . Dans le même ordre d'idée, si  $\text{Span}\{\phi_{\alpha_i}\}$  est dense, alors en prenant n'importe quelle queue de suite, ça reste dense.

Prouvons donc l'énoncé (17.57). Si  $\text{Span}\{1, \phi_{\alpha_i}\}$  est dense, alors en posant  $\beta_1 = 0$ ,  $\beta_i = \alpha_{i-1}$  notre théorème prouve que  $\sum_{i=2}^{\infty} \frac{1}{\beta_i} = +\infty$ , cela est exactement que  $\sum_{i=1}^{\infty} \frac{1}{\alpha_i} = +\infty$ . Dans l'autre sens, si  $\sum_{i \geq 1} \frac{1}{\alpha_i} = +\infty$ , alors nous avons aussi  $\sum_{i \geq 2} \frac{1}{\alpha_i} = +\infty$  et notre théorème dit que  $\text{Span}\{\phi_{\alpha_i}\}$  est dense. A fortiori,  $\text{Span}\{1, \phi_{\alpha_i}\}$  est dense.

**Exemple 17.9.**

Nous savons depuis le théorème 15.129 que la somme des inverses des nombres premiers diverge.  $\triangle$

**17.3 Intégrales convergeant uniformément****17.3.1 Définition et propriété**

DEFooSHWAooWtswtp

**Définition 17.10.**

Soit  $(\Omega, \mu)$  un espace mesuré. Nous disons que l'intégrale

$$\int_{\Omega} f(x, \omega) d\mu(\omega) \quad (17.58)$$

**converge uniformément** en  $x$  si pour tout  $\epsilon > 0$ , il existe un compact  $K_{\epsilon}$  tel que pour tout compact  $K$  tel que  $K_{\epsilon} \subset K$  nous avons

$$\left| \int_{\Omega \setminus K} f(x, \omega) d\mu(\omega) \right| \leq \epsilon. \quad (17.59)$$

Le point important est que le choix de  $K_{\epsilon}$  ne dépend pas de  $x$ .

LemOgQdpJ

**Lemme 17.11.**

Soit

$$F(x) = \int_{\Omega} f(x, \omega) d\mu(\omega), \quad (17.60)$$

une intégrale uniformément convergente. Pour chaque  $k \in \mathbb{N}$  nous considérons un compact  $K_k$  tel que

$$\left| \int_{\Omega \setminus K_k} f(x, \omega) d\mu(\omega) \right| \leq \frac{1}{k}. \quad (17.61)$$

Alors la suite de fonctions  $F_k$  définie par

$$F_k(x) = \int_{K_k} f(x, \omega) d\mu(\omega) \quad (17.62)$$

converge uniformément vers  $F$ .

---

8. Dont le rapport du jury 2014

*Démonstration.* Nous avons

$$|F_k(x) - F(x)| = \left| \int_{K_k} f(x, \omega) d\mu(\omega) - \int_{\Omega} f(x, \omega) d\mu(\omega) \right| \quad (17.63a)$$

$$= \left| \int_{\Omega \setminus K_k} f(x, \omega) d\mu(\omega) \right| \quad (17.63b)$$

$$\leq \frac{1}{k}. \quad (17.63c)$$

□

### 17.3.2 Critères de convergence uniforme

Afin de tester l'uniforme convergence d'une intégrale, nous avons le **critère de Weierstrass** :

ThoCritWeiIntUnifCv

#### Théorème 17.12.

Soit  $f(x, t) : [\alpha, \beta] \times [a, \infty[ \rightarrow \mathbb{R}$ , une fonction dont la restriction à toute demi-droite  $x = c$  est mesurable. Si  $|f(x, t)| < \varphi(t)$  et  $\int_a^\infty \varphi(t) dt$  existe, alors l'intégrale

$$\int_0^\infty f(x, t) dt \quad (17.64)$$

est uniformément convergente.

Le théorème suivant est le **critère d'Abel** :

ThoAbelIntUnif

#### Théorème 17.13.

Supposons que  $f(x, t) = \varphi(x, t)\psi(x, t)$  où  $\varphi$  et  $\psi$  sont bornées et intégrables en  $t$  au sens de Riemann sur tout compact  $[a, b]$ ,  $b \geq a$ . Supposons que :

- (1)  $\left| \int_a^T \varphi(x, t) dt \right| \leq M$  où  $M$  est indépendant de  $T$  et de  $x$ ,
- (2)  $\psi(x, t) \geq 0$ ,
- (3) pour tout  $x \in [\alpha, \beta]$ ,  $\psi(x, t)$  est une fonction décroissante de  $t$ ,
- (4) les fonctions  $x \mapsto \psi(x, t)$  convergent uniformément vers 0 lorsque  $t \rightarrow \infty$ .

Alors l'intégrale

$$\int_a^\infty f(x, t) dt \quad (17.65)$$

est uniformément convergente.

#### Remarque 17.14.

Étant donné que la fonction sinus est bornée, il est tentant de l'utiliser comme  $\varphi$  dans le critère d'Abel (théorème 17.13). Hélas,

$$\int_0^T \sin(xt) dt = -\frac{1}{x} (\cos(xT) - \cos(x)), \quad (17.66)$$

qui n'est pas bornée en  $x$ . Poser  $\varphi(x, t) = \sin(xt)$  ne fonctionne pas pour assurer la convergence uniforme sur un intervalle qui contient des  $x$  arbitrairement proches de 0. Le critère d'Abel avec  $\varphi(x, t) = \sin(xt)$  ne permet que de conclure à l'uniforme convergence *sur tout compact* ne contenant pas 0. C'est toutefois souvent suffisant pour étudier la continuité ou la dérivabilité en se servant coup du compact, voir 17.18.

## 17.4 Fonctions définies par une intégrale

SecCHwnBDj

Soit  $(\Omega, \mu)$  un espace mesuré. Nous nous demandons dans quel cas l'intégrale

$$F(x) = \int_{\Omega} f(x, \omega) d\omega \quad (17.67)$$

définit une fonction  $F$  continue, dérivable ou autre.

Dans la suite nous allons considérer des fonctions  $f$  à valeurs réelles. Quitte à passer aux composantes, nous pouvons considérer des fonctions à valeurs vectorielles. Par contre le fait que  $x$  soit dans  $\mathbb{R}$  ou dans  $\mathbb{R}^n$  n'est pas spécialement une chose facile à traiter.

### 17.4.1 Continuité sous l'intégrale

Nous allons présenter deux théorèmes donnant la continuité de  $F$ .

- (1) Si  $f$  est majorée par une fonction ne dépendant pas de  $x$ , nous avons le théorème 17.15,
- (2) si l'intégrale est uniformément convergente, nous avons le théorème 17.16.

ThoKnuSNd

#### Théorème 17.15.

Soit  $(\Omega, \mu)$  est un espace mesuré, soit  $x_0 \in \mathbb{R}^m$  et  $f: U \times \Omega \rightarrow \mathbb{R}$  où  $U$  est ouvert dans  $\mathbb{R}^m$ . Nous supposons que

- (1) Pour chaque  $x \in \mathbb{R}^m$ , la fonction  $\omega \mapsto f(x, \omega)$  est dans  $L^1(\Omega, \mu)$ .
- (2) Pour chaque  $\omega \in \Omega$ , la fonction  $x \mapsto f(x, \omega)$  est continue en  $x_0$ .
- (3) Il existe une fonction  $G \in L^1(\Omega)$  telle que

ItemNAuYNG

$$|f(x, \omega)| \leq G(\omega) \quad (17.68)$$

pour tout  $x \in U$ .

Alors la fonction

$$\begin{aligned} F: U &\rightarrow \mathbb{R} \\ x &\mapsto \int_{\Omega} f(x, \omega) d\mu(\omega) \end{aligned} \quad (17.69)$$

est continue en  $x_0$ .

*Démonstration.* Soit  $(x_n)$  une suite convergente vers  $x_0$ . Nous considérons la suite de fonctions  $f_n: \Omega \rightarrow \mathbb{R}$  définies par

$$f_n(\omega) = f(x_n, \omega). \quad (17.70)$$

sur qui nous pouvons utiliser le théorème de la convergence dominée (théorème 14.207) pour obtenir

$$\lim_{n \rightarrow \infty} F(x_n) = \lim_{n \rightarrow \infty} \int_{\Omega} f(x_n, \omega) d\mu(\omega) \quad (17.71a)$$

$$= \int_{\Omega} \lim_{n \rightarrow \infty} f(x_n, \omega) d\mu(\omega) \quad (17.71b)$$

$$= \int_{\Omega} f(x_0, \omega) d\mu(\omega) \quad (17.71c)$$

$$= F(x_0). \quad (17.71d)$$

Nous avons utilisé la continuité de  $f(., \omega)$ . □

Si nous avons un peu de compatibilité entre la topologie et la mesure, alors nous pouvons utiliser l'uniforme convergence d'une intégrale pour obtenir la continuité d'une fonction définie par une intégrale.

ThotexmgE

#### Théorème 17.16.

Soit  $(\Omega, \mu)$  un espace topologique mesuré tel que tout compact est de mesure finie. Soit une fonction  $f: \mathbb{R} \times \Omega \rightarrow \mathbb{R}$  telle que

- (1) Pour chaque  $x \in \mathbb{R}$ , la fonction  $f(x, .)$  est  $L^1(\Omega, \mu)$ .
- (2) Pour chaque  $\omega \in \Omega$ , la fonction  $f(., \omega)$  est continue en  $x_0$ .

(3) *L'intégrale*

$$F(x) = \int_{\Omega} f(x, \omega) d\mu(\omega) \quad (17.72)$$

*est uniformément convergente*<sup>9</sup>.

Alors la fonction  $F$  est continue en  $x_0$ .

*Démonstration.* Nous reprenons les notations du lemme 17.11. Les fonctions

$$F_k(x) = \int_{K_k} f(x, \omega) d\mu(\omega) \quad (17.73)$$

existent parce que les fonctions  $f(x, \cdot)$  sont dans  $L^1(\Omega)$ . Montrons que les fonctions  $F_k$  sont continues. Soit une suite  $x_k \rightarrow x_0$  nous avons

$$\lim_{n \rightarrow \infty} F_k(x_n) = \lim_{n \rightarrow \infty} \int_{K_k} f(x_n, \omega) d\mu(\omega). \quad (17.74)$$

Nous pouvons inverser la limite et l'intégrale en utilisant le théorème de la convergence dominée. Pour cela, la fonction  $f(x_n, \omega)$  étant continue sur le compact  $K_k$ , elle y est majorée par une constante. Le fait que les compacts soient de mesure finie (hypothèse) implique que les constantes soient intégrables sur  $K_k$ . Le théorème de la convergence dominée implique alors que

$$\lim_{n \rightarrow \infty} F_k(x_n) = \int_{K_k} \lim_{n \rightarrow \infty} f(x_n, \omega) d\mu(\omega) = \int_{K_k} f(x_0, \omega) d\mu(\omega) = F_k(x_0). \quad (17.75)$$

Nous avons utilisé le fait que  $f(\cdot, \omega)$  était continue en  $x_0$ .

Le lemme 17.11 nous indique alors que la convergence  $F_k \rightarrow F$  est uniforme. Les fonctions  $F_k$  étant continues, la fonction  $F$  est continue.  $\square$

Pour finir, citons ce résultat concernant les fonctions réelles.

ThoInDerrtCvUniffContinue

### **Théorème 17.17.**

Nous considérons  $F(x) = \int_a^\infty f(x, t) dt$ . Si  $f$  est continue sur  $[\alpha, \beta] \times [a, \alpha[$  et l'intégrale converge uniformément, alors  $F(x)$  est continue.

## **17.4.2 Le coup du compact**

NORMMooZWECoohvRgBw

### **17.18.**

Nous avons vu des fonctions définies par toute une série de processus de limite (suites, séries, intégrales). Une des questions centrales est de savoir si la fonction limite est continue, dérivable, intégrable, etc. étant donné que les fonctions sont continues.

Pour cela, nous inventons le concept de *convergence uniforme*. Si la limite (série, intégrale) est uniforme, alors la fonction limite sera continue. Il arrive qu'une limite ne soit pas uniforme sur un intervalle ouvert  $]0, 1]$ , et que nous voulions quand même prouver la continuité sur cet intervalle. C'est à cela que sert la notion de convergence uniforme *sur tout compact*. En effet, la notion de continuité est une notion locale : savoir ce qu'il se passe dans un petit voisinage autour de  $x$  est suffisant pour savoir la continuité en  $x$  (idem pour sa dérivée).

Si nous avons convergence uniforme sur tout compact de  $]0, 1]$ , mais pas convergence uniforme sur cet intervalle, la limite sera quand même continue sur  $]0, 1]$ . En effet, si  $x \in ]0, 1]$ , il existe un ouvert autour de  $x$  contenu dans un compact contenu dans  $]0, 1]$ . L'uniforme convergence sur ce compact suffit à prouver la continuité en  $x$ .

Déduire la continuité sur un ouvert à partir de l'uniforme convergence sur tout compact de l'ouvert est appelé faire le **coup du compact**.

9. Définition 17.10.

### 17.4.3 Dérivabilité sous l'intégrale

Nous traitons à présent de la dérivabilité de la fonction  $F$  définie comme intégrale de  $f$ . Dans le théorème 17.19 nous traitons de fonctions sur  $\mathbb{R}$  à valeurs dans  $\mathbb{C}$ . Pour les fonctions définies sur  $\mathbb{C}$ , voir le théorème 26.73.

**Théorème 17.19** (Dérivation sous le signe intégral, formule de Leibniz, thème 62[523, 559]).

Soit  $(\Omega, \mu)$  un espace mesuré, un intervalle ouvert  $I \subset \mathbb{R}$ , et une fonction  $f: \mathbb{R} \times \Omega \rightarrow \mathbb{C}$ . Nous supposons qu'il existe  $A$  mesurable de mesure nulle dans  $\Omega$  tels que

(1) Pour tout  $x \in I$ , la fonction  $\omega \mapsto f(x, \omega)$  soit dans  $L^1(\Omega)$ .

(2) L'application  $x \mapsto f(x, \omega)$  est dérivable<sup>10</sup> pour tout  $x \in I$  et pour tout  $\omega \in \Omega \setminus A$ .

(3) Il existe une fonction  $G: \Omega \rightarrow \mathbb{R}^+$  intégrable sur  $\Omega$  telle que

$$\left| \frac{\partial f}{\partial x}(x, \omega) \right| \leq G(\omega) \quad (17.76)$$

pour tout  $x \in I$  et pour tout  $\omega \in \Omega \setminus A$ .

Alors la fonction

$$\begin{aligned} F: \mathbb{R} &\rightarrow \mathbb{C} \\ x &\mapsto \int_{\Omega} f(x, \omega) d\mu(\omega) \end{aligned} \quad (17.77)$$

est dérivable sur  $I$  et pour tout  $a \in I$  nous avons

$$F'(a) = \int_{\Omega} \frac{\partial f}{\partial x}(a, \omega) d\mu(\omega), \quad (17.78)$$

*Démonstration.* Soit une suite  $(x_n)$  dans  $I$  telle que  $x_n \neq a$  et  $x_n \rightarrow a$ . Si la limite

$$\lim_{n \rightarrow \infty} \frac{F(a) - F(x_n)}{a - x_n} \quad (17.79)$$

existe et ne dépend pas de la suite choisie, alors la fonction  $F$  est dérivable en  $a$  et sa dérivée vaut cette limite. Autrement dit, nous nous mettons en devoir d'étudier la limite

$$\lim_{n \rightarrow \infty} \int_{\Omega} \frac{f(a, \omega) - f(x_n, \omega)}{a - x_n} d\omega. \quad (17.80)$$

montrer qu'elle existe, ne dépend pas de la suite choisie et vaut  $\int_{\Omega} \partial_x f(a, \omega) d\omega$ . On y va.

(1) **La bonne suite de fonctions** D'abord nous posons

$$g_n(\omega) = \frac{f(x_n, \omega) - f(a, \omega)}{x_n - a}. \quad (17.81)$$

Nous montrons à présent que cette suite vérifie les hypothèses du théorème de la convergence dominée 14.207.

- Chacune des fonctions  $g_n$  est dans  $L^1(\Omega)$  parce que,  $a$  étant fixé, l'élément  $x_n$  est dans  $I \setminus \{a\}$ ; le dénominateur n'a donc aucun rôle. L'hypothèse (1) montre que  $\omega \mapsto f(x_n, \omega)$  et  $\omega \mapsto f(a, \omega)$  sont dans  $L^1$ . La somme est donc dans  $L^1$  (proposition 14.193).
- Par l'hypothèse (2), pour chaque  $\omega$  nous avons une fonction dérivable. Nous pouvons donc passer à la limite :

$$\lim_{n \rightarrow \infty} g_n(\omega) = \frac{\partial f}{\partial x}(a, \omega). \quad (17.82)$$

<sup>10</sup>. La dérivabilité pour une fonction à valeurs dans  $\mathbb{C}$  n'a rien de mystérieux : c'est la dérivée composante par composante. Rien à voir avec la dérivée complexe.

- En ce qui concerne la majoration de  $g_n$ , nous utilisons le théorème des accroissements finis [12.216\(2\)](#). Pour chaque  $\omega$ , ce théorème peut être utilisé sur la fonction  $x \mapsto f(x, \omega)$  qui est dérivable. Nous avons :

$$\left| \frac{f(x_n, \omega) - f(a, \omega)}{x_n - a} \right| \leq \sup_{x \in [a, x_n]} \left| \frac{\partial f}{\partial x}(x, \omega) \right| \leq G(\omega). \quad (17.83)$$

Nous avons utilisé l'hypothèse [\(3\)](#).

Les hypothèses de la convergence dominée sont satisfaites.

- (2) **Convergence dominée** Le théorème de la convergence dominée de Lebesgue (théorème [14.207](#)) nous permet alors de calculer la limite [\(17.80\)](#) :

$$\lim_{n \rightarrow \infty} \int_{\Omega} g_n(\omega) d\omega = \int_{\Omega} \lim_{n \rightarrow \infty} g_n(\omega) d\omega = \int_{\Omega} \frac{\partial f}{\partial x}(a, \omega) d\omega. \quad (17.84)$$

Notons que l'existence de la dernière intégrale fait partie du théorème de la convergence dominée.

Nous avons donc prouvé que la limite de gauche existait et ne dépendait pas de la suite choisie. Donc  $F$  est dérivable en  $a$  et la dérivée vaut cette limite :

$$F'(a) = \int_{\Omega} \frac{\partial f}{\partial x}(a, \omega) d\mu(\omega). \quad (17.85)$$

□

En ce qui concerne les fonctions dans  $\mathbb{R}^n$ , il y a les propositions [17.24](#) et [17.25](#) qui parlent de différentiabilité sous l'intégrale.

#### ⚠ Avertissement/question au lecteur !! [17.20](#)

Attention : l'énoncé et la démonstration de la proposition [17.21](#) sont de moi. Écrivez-moi pour

- me dire si ça vous semble correct (ou pas),
- me donner un lien vers une source qui énonce et démontre ce résultat.

PROPOoJKXJooLxgEGd

#### Proposition 17.21 ([\[1\]](#)).

Soient un espace mesuré  $(\Omega, \mu)$  ainsi qu'une fonction  $f: \mathbb{R}^n \times \Omega \rightarrow \mathbb{C}$ . Si  $\omega \in \Omega$  est fixé, nous notons

$$\begin{aligned} f_{\omega}: \mathbb{R}^n &\rightarrow \mathbb{C} \\ x &\mapsto f(x, \omega) \end{aligned} \quad (17.86)$$

et si  $x \in \mathbb{R}^n$  est fixé, nous notons

$$\begin{aligned} f_x: \Omega &\rightarrow \mathbb{C} \\ \omega &\mapsto f(x, \omega). \end{aligned} \quad (17.87)$$

Soient  $\delta > 0$ ,  $A$  de mesure nulle dans  $\Omega$  et une liste d'indices<sup>[11](#)</sup>  $\alpha$  tels que

- (1) pour tout  $x \in \mathbb{R}^n$ , la fonction  $f_x$  est dans  $L^1(\Omega)$ ,
- (2) la dérivée partielle multiple  $\partial^{\alpha} f_{\omega}(x)$  existe pour tout  $x \in B(a, \delta)$  et pour tout  $\omega \in A^c$ .
- (3) pour toute queue  $\beta$  de la liste d'indice  $\alpha$ , il existe une fonction  $G_{\beta} \in L^1(\Omega)$  telle que

$$|(\partial^{\beta} f_{\omega})(x)| \leq G_{\beta}(\omega) \quad (17.88)$$

pour tout  $x \in B(a, \delta)$  et  $\omega \in A^c$ .

Enfin nous posons

$$F(x) = \int_{\Omega} f(x, \omega) d\mu(\omega). \quad (17.89)$$

Alors pour toute queue  $\beta$  de suite  $\alpha$ , nous avons

---

11. Voir [-1.1](#).

- (1)  $(\partial^\beta F)$  existe en  $a$ ,  
 (2) nous avons la formule

$$(\partial^\beta F)(x) = \int_{\Omega} (\partial^\beta f_\omega)(a) d\mu(\omega). \quad (17.90)$$

*Démonstration.* Nous allons opérer une récurrence sur  $\alpha$ . Plus précisément, si  $|\alpha| = p$ , nous allons ajouter une dérivation à la fois, et dans l'ordre inverse de  $\alpha$ . Donc nous commençons par  $\partial_{\alpha_p}$ , puis par  $(\partial_{p-1}, \partial_p)$ , etc.

Nous commençons par prouver la formule (17.90) dans le cas de  $\beta = (\alpha_p)$ . Et pour alléger les notations nous notons  $\alpha_p = i$ . Nous posons

$$\begin{aligned} \varphi: \mathbb{R} \times \Omega &\rightarrow \mathbb{C} \\ (t, \omega) &\mapsto f(a + te_i, \omega). \end{aligned} \quad (17.91)$$

Posons  $H(t) = F(a + te_i)$ , c'est-à-dire

$$H(t) = \int_{\Omega} f(a + te_i, \omega) d\mu(\omega) = \int_{\Omega} \varphi(t, \omega) d\mu(\omega). \quad (17.92)$$

En utilisant le théorème 17.19 sur la fonction  $\varphi$  nous trouvons

$$H'(0) = \int_{\Omega} \varphi'(0, \omega) d\mu(\omega) = \int_{\Omega} \frac{\partial f}{\partial x_i}(a, \omega) d\mu(\omega), \quad (17.93)$$

ce qui est la formule demandée dans le cas  $\alpha = (i)$ .

Pour la récurrence, nous supposons que la formule est démontrée pour  $\beta = (\alpha_{p-k}, \dots, \alpha_p)$ , et nous montrons qu'elle fonctionne encore pour  $\sigma = (i, \beta)$ .

Il s'agit simplement de remarquer que la fonction

$$g(x, \omega) = (\partial^\beta f_\omega)(x) \quad (17.94)$$

vérifie encore les conditions du théorème 17.19<sup>12</sup>. □

#### 17.4.4 Dérivation et intégrale

##### ⚠ Avertissement/question à la lectrice !! 17.22

Si je regarde bien, le théorème 17.23 n'est utilisé nulle part. Je vais donc le supprimer un jour. Si vous l'aimez, vous devriez m'en envoyer une démonstration.

##### Théorème 17.23.

Soient  $A$  un intervalle ouvert de  $\mathbb{R}$  et  $\Omega$  un espace mesuré. Soient une fonction  $f: A \times \Omega \rightarrow \mathbb{R}$  et

$$F(x) = \int_{\Omega} f(x, \omega) d\omega. \quad (17.95)$$

Nous supposons les points suivants.

- (1) La fonction  $f$  est mesurable.
- (2) Pour chaque  $x \in A$ , la fonction  $\omega \mapsto f(x, \omega)$  est intégrable sur  $\Omega$ .
- (3) Pour presque tout  $\omega \in \Omega$ , la fonction  $f(x, \omega)$  est une fonction absolument continue<sup>13</sup> de  $x$ .
- (4) La fonction  $\frac{\partial f}{\partial x}$  est localement intégrable, c'est-à-dire que pour tout  $[a, b] \subset A$ ,

$$\int_a^b \int_{\Omega} \left| \frac{\partial f}{\partial x}(x, \omega) \right| d\omega dx < \infty. \quad (17.96)$$

Alors la fonction  $F$  est absolument continue et pour presque tout  $x \in A$ , la dérivée est donné par

$$\frac{d}{dx} \int_{\Omega} f(x, \omega) d\omega = \int_{\Omega} \frac{\partial f}{\partial x}(x, \omega) d\omega. \quad (17.97)$$

*Démonstration.* Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

12. Je n'ai pas fait cette vérification. Écrivez-moi si vous l'avez faite.

13. Définition 14.314 et proposition 14.317.

### 17.4.5 Différentiabilité sous l'intégrale

Le théorème suivant est restrictif sur l'ensemble d'intégration (qui doit être compact), mais accepte des fonctions de plusieurs variables, ce qui est un premier pas vers la différentiabilité.

PropDerrSSIntegraleDSD

**Proposition 17.24** (Dérivation sous l'intégrale).

Supposons  $A \subset \mathbb{R}^m$  ouvert et  $B \subset \mathbb{R}^n$  compact. Nous considérons une fonction  $f: A \times B \rightarrow \mathbb{R}$ . Si pour un  $i \in \{1, \dots, n\}$ , la dérivée partielle  $\frac{\partial f}{\partial x_i}$  existe dans  $A \times B$  et est continue, alors la fonction

$$F(x) = \int_B f(x, t) dt \quad (17.98)$$

admet une dérivée partielle dans la direction  $x_i$  sur  $A$ . Cette dérivée partielle  $y$  est continue et

$$\frac{\partial F}{\partial x_i}(a) = \int_B \frac{\partial f}{\partial x_i}(a, t) dt, \quad (17.99)$$

pour tout  $a$  dans l'ouvert  $A$ .

*Démonstration.* Nous procédons en plusieurs étapes.

(1) **F est dérivable** Nous voulons prouver que  $\frac{\partial F}{\partial x_i}(a, t)$  existe. Pour cela nous posons

$$g_l(t) = \frac{f(a_1, \dots, a_i + \epsilon_l, \dots, a_n, t) - f(a_1, \dots, a_i, \dots, a_n, t)}{\epsilon_l} \quad (17.100)$$

où  $\epsilon_l$  est une suite de nombres tendant vers zéro. La fonction  $f$  est dérivable dans la direction  $x_i$  si et seulement si  $\lim_{l \rightarrow \infty} g_l(t)$  existe et ne dépend pas du choix de la suite. À ce moment, la valeur de la dérivée partielle sera cette limite. Dans notre cas, nous savons que  $f$  admet une dérivée partielle dans la direction  $x_i$  et donc nous avons

$$\frac{\partial f}{\partial x_i}(a, t) = \lim_{l \rightarrow \infty} g_l(t). \quad (17.101)$$

De la même façon pour  $F$  nous avons

$$\frac{\partial F}{\partial x_i} = \lim_{l \rightarrow \infty} \int_B g_l(t) dt. \quad (17.102)$$

Sous-entendu : si la limite de droite ne dépend pas de la suite choisie, alors  $\frac{\partial F}{\partial x_i}$  existe et vaut cette limite.

Considérant la continuité de  $f$ , le seul point à vérifier pour le théorème de la convergence dominée de Lebesgue est l'existence d'une fonction intégrable de  $t$  majorant  $g_l$ . Pour cela le théorème de accroissements finis (théorème 12.216) appliqué à la fonction  $\epsilon \mapsto f(a_n, \dots, a_i + \epsilon, \dots, a_n)$  nous dit que

$$f(a_1, \dots, a_i + \epsilon_l, \dots, a_n, t) - f(a_1, \dots, a_i, \dots, a_n, t) = \epsilon_l \frac{\partial f}{\partial x_i}(a_1, \dots, \theta, \dots, a_n, t) \quad (17.103)$$

pour un certain  $\theta \in B(a_i, \epsilon_l)$ . Notons que ce  $\theta$  dépend de  $t$  mais pas de  $l$ . Vu que  $\partial_i f$  est continue par rapport à ses deux variables, si  $K$  est un voisinage compact autour de  $a$ , il existe  $M > 0$  tel que

$$\left| \frac{\partial f}{\partial x_i}(x, t) \right| < M \quad (17.104)$$

EqMXqviPC

pour tout  $x \in K$  et tout  $t \in B$ . La valeur de  $\frac{\partial f}{\partial x_i}(a_1, \dots, \theta, \dots, a_n, t)$  est donc bien majorée par rapport à  $\theta$  et par rapport à  $t$  en même temps par une constante qui n'a pas de mal à être intégrée sur le compact  $B$ .



Le théorème de la convergence dominée (théorème 14.207) s'applique donc bien et nous avons

$$\lim_{l \rightarrow \infty} \int_B g_l(t) dt = \int_B \lim_{l \rightarrow \infty} g_l(t) = \int_B \frac{\partial f}{\partial x_i}(a, t) dt. \quad (17.105)$$

Le membre de droite ne dépendant pas de la suite  $\epsilon_l$  choisie, le membre de gauche est bien la dérivée de  $F$  par rapport à  $x_i$  et nous avons

$$\frac{\partial F}{\partial x_i}(a) = \int_B \frac{\partial f}{\partial x_i}(a, t) dt. \quad (17.106)$$

Cela prouve la première partie de la proposition.

- (2) **La dérivée est continue** Soit  $K$  un voisinage compact autour de  $a$  et  $U'$  un ouvert tel que  $a \in U' \subset K$ . Nous avons encore la majoration (17.104) sur  $U'$  et donc le théorème de continuité sous l'intégrale 17.15 nous indique que la fonction

$$\begin{aligned} U' &\rightarrow \mathbb{R} \\ x &\mapsto \int_B \frac{\partial f}{\partial x_i}(x, t) dt \end{aligned} \quad (17.107)$$

est continue en  $a$ . □

Une conséquence de la proposition 17.24 est que si elle fonctionne pour tous les  $i$ , alors  $F$  est différentiable et même de classe  $C^1$ , et la différentielle de  $F$  s'obtient comme intégrale de la différentielle de  $f$ .

PropAOZkDsh

**Proposition 17.25.**

Supposons  $A \subset \mathbb{R}^m$  ouvert et  $B \subset \mathbb{R}^n$  compact. Si pour tout  $i \in \{1, \dots, n\}$ , la dérivée partielle  $\frac{\partial f}{\partial x_i}$  existe dans  $A \times B$  et est continue, alors  $F$  est de classe  $C^1$  et

$$(dF)_a = \int_B (df_t)_a dt \quad (17.108)$$

où  $f_t(x) = f(x, t)$ .

*Démonstration.* En vertu de la proposition 17.24, toutes les dérivées partielles de  $F$  sont continues. Cela implique que  $F$  est de classe  $C^1$  par le théorème 12.320 et que la différentielle s'écrit en termes des dérivées partielles avec la formule usuelle. Nous avons alors

$$(dF)_a(u) = \sum_k \frac{\partial F}{\partial x_k}(a) u_k \quad (17.109a)$$

$$= \int_B \sum_k \frac{\partial f}{\partial x_k}(a, t) dt \quad (17.109b)$$

$$= \int_B \sum_k \frac{\partial f_t}{\partial x_k}(a) u_k dt \quad (17.109c)$$

$$= \int_B (df_t)_a(u) dt. \quad (17.109d)$$

Ce qui est la formule annoncée. □

Un autre théorème tourne autour du pot, et me semble inutile.

ThoOLAQyRL

**Théorème 17.26.**

Soit  $(\Omega, \mu)$  un espace mesuré, une fonction  $f: \mathbb{R}^n \times \Omega \rightarrow \mathbb{R}$  et  $a \in \mathbb{R}^n$ . Nous considérons la fonction

$$F(x) = \int_{\Omega} f(x, \omega) d\mu(\omega). \quad (17.110)$$

Pour chaque  $k = 1, \dots, n$  nous supposons avoir

$$\frac{\partial F}{\partial x_k}(a) = F'_{|_k}(a) = \int_{\Omega} \frac{\partial f_{|_k}}{\partial t}(a_k, \omega) d\mu(\omega) \quad (17.111)$$

où  $F_{|_k}(t) = F(a_1, \dots, t, \dots, a_n)$  et  $f_{|_k}$  est définie de façon similaire.

Nous supposons de plus que les fonctions  $\partial_{x_k} F$  sont continues.

Alors  $F$  est de classe  $C^1$  et sa différentielle est donnée par

$$dF_a = \int_{\Omega} (df_{\omega})_a d\omega \quad (17.112)$$

où  $f_{\omega}$  est définie par  $f_{\omega}(x) = f(x, \omega)$ .

*Démonstration.* Étant donné que les dérivées partielles de  $F$  en  $a$  existent et sont continues, le théorème 12.320 dit que  $F$  est différentiable et que

$$dF_a(u) = \sum_{k=1}^n \frac{\partial F}{\partial x_k}(a) u_k. \quad (17.113)$$

La linéarité de l'intégrale et les hypothèses nous donnent alors

$$dF_a(u) = \sum_{k=1}^n \frac{\partial F}{\partial x_k}(a) u_k \quad (17.114a)$$

$$= \int_{\Omega} \sum_k \frac{\partial f_{|_k}}{\partial t}(a_k; \omega) u_k d\mu(\omega) \quad (17.114b)$$

$$= \int_{\Omega} \sum_k \frac{\partial f}{\partial x_k}(a; \omega) u_k d\mu(\omega) \quad (17.114c)$$

$$= \int_{\Omega} (df_{\omega})_a(u) d\mu(\omega), \quad (17.114d)$$

et donc  $dF_a = \int_{\Omega} (df_{\omega})_a d\mu(\omega)$ . □

Notons qu'en passant aux composantes, ce théorème fonctionne tout aussi bien pour des fonctions à valeurs dans un espace vectoriel normé de dimension finie plutôt que dans  $\mathbb{R}$ .

## 17.5 Théorème d'Hadamard

LemWNB00GP1IwT

**Lemme 17.27** (Hadamard[560]).

Soit une fonction  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^p$  avec  $p \geq 1$ . Pour tout  $a \in \mathbb{R}^n$  il existe des fonctions  $g_1, \dots, g_n$  de classe  $C^{p-1}$  telles que

$$f(x) = f(a) + \sum_{i=1}^n (x_i - a_i) g_i(x). \quad (17.115)$$

*Démonstration.* Puisque  $f$  est de classe  $C^1$ , le théorème fondamental de l'analyse 14.288 s'applique et

$$f(x) - f(a) = \int_0^1 \frac{d}{dt} [f(a + t(x - a))] dt = \int_0^1 \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a + t(x - a)) (x_i - a_i) dt. \quad (17.116) \quad \text{EqZLT00VKmCln}$$

Plus de détails : la fonction  $t \mapsto \frac{d}{dt} [f(a + t(x - a))]$  possède comme primitive la fonction  $F(t) = f(a + t(x - a))$ .

Nous posons

$$g_i(x) = \int_0^1 \frac{\partial f}{\partial x_i}(a + t(x - a)) dt \quad (17.117)$$

L'intégrale existe parce qu'il s'agit d'une fonction continue sur un compact et donc majorée par une constante. Pour voir que  $g_i$  est de classe  $C^{p-1}$  nous pouvons calculer  $\frac{\partial g_i}{\partial x_k}$  en permutant dérivée et intégrale par la proposition 17.24 :

$$\frac{\partial g_i}{\partial x_k}(x) = \int_0^1 \frac{\partial}{\partial x_k} \left( \frac{\partial f}{\partial x_i}(a + t(x-a)) \right) dt = \int_0^1 t \frac{\partial^2 f}{\partial x_k \partial x_i}(a + t(x-a)) dt. \quad (17.118)$$

Nous pouvons ainsi permuter  $p-1$  dérivées tout en gardant une fonction continue dans l'intégrale. Le théorème 17.15 nous donne alors une fonction continue. Ainsi toutes les fonctions

$$\frac{\partial^{p-1} g_i}{\partial x_{i_1} \dots \partial x_{i_{p-1}}} \quad (17.119)$$

sont continues et  $g_i$  est de classe  $C^{p-1}$  par le théorème 12.356.

En repartant de (17.116) nous avons alors bien ce qui était annoncé :

$$f(x) = f(a) + \sum_{i=1}^n g_i(x)(x_i - a_i). \quad (17.120)$$

□

CorQBXHooZVKeNG

### Corolaire 17.28.

Soit  $\phi \in \mathcal{D}(\mathbb{R})$  tel que  $\phi^{(k)}(x_0) = 0$  pour tout  $k \leq n$ . Alors il existe une fonction  $\psi \in \mathcal{D}(\mathbb{R})$  telle que

$$\phi(x) = (x - x_0)^{n+1} \psi(x) \quad (17.121)$$

pour tout  $x \in \mathbb{R}$ .

*Démonstration.* En utilisant le lemme de Hadamard 17.27 avec  $a = x_0$ ,  $n = 1$  et  $f(x_0) = 0$ , nous avons une fonction  $g_1$  à support compact telle que

$$\phi(x) = \phi(x_0) + (x - x_0)g_1(x). \quad (17.122)$$

EqTOJGooWZBBRJ

Alors  $\phi'(x) = g_1(x) + (x - x_0)g_1'(x)$ , ce qui donne immédiatement  $g_1(x_0) = 0$  et donc une fonction  $g_2$  telle que  $g_1(x) = (x - x_0)g_2(x)$ . En injectant dans (17.122) nous avons

$$\phi(x) = (x - x_0)^2 g_2(x). \quad (17.123)$$

Il suffit de continuer ainsi tant que les dérivées de  $\phi$  s'annulent. □

## 17.6 Deux théorèmes de point fixe

Nous allons voir Picard. Les autres théorème de point fixe que sont Brouwer, Schauder et Markov-Kakutani sont plus bas<sup>14</sup> parce qu'ils utilisent de l'intégration. Voir le thème 65 pour les retrouver.

### 17.6.1 Points fixes attractifs et répulsifs

DEFooTMZUooMoBDGC

#### Définition 17.29 ([1]).

Soit  $I$  un intervalle fermé de  $\mathbb{R}$  et  $\varphi: I \rightarrow I$  une application  $C^1$ . Soit  $a$  un point fixe de  $\varphi$ . Nous disons que  $a$  est **attractif** si il existe un voisinage  $V$  de  $a$  tel que pour tout  $x_0 \in V$  la suite  $x_{n+1} = \varphi(x_n)$  converge vers  $a$ .

Le point  $a$  est dit **répulsif** si il existe un voisinage<sup>15</sup>  $V$  de  $a$  tel que pour toute suite de la forme  $x_{n+1} = \varphi(x_n)$ , si  $u_n \in V$ , alors  $|x_{n+1} - a| \geq |u_n - a|$ .

14. Dans la section 20.5.

15. Cette partie de la définition est de moi. Je suis loin d'être sûr que c'est une bonne définition. Si vous avez une idée, écrivez-moi. Sinon, tenez-vous en à prendre le lemme 17.31 comme définition.

**17.30.**

Beaucoup de sources prennent le lemme 17.31 comme définition de point fixe attractif et répulsif. Celle en termes de suites ne demande au moins pas la dérivabilité de  $f$ . Elle est donc un tout petit peu plus générale.

LEMooTIRAooQIHLYt

**Lemme 17.31** ([561, 562]).

Soient un intervalle  $I$  de  $\mathbb{R}$ , et une application  $\varphi: I \rightarrow I$  de classe  $C^1$ . Soit  $a$  un point fixe de  $\varphi$ .

- (1) Si  $|\varphi'(a)| < 1$  alors  $a$  est attractif et la convergence est au moins exponentielle.
- (2) Si  $|\varphi'(a)| > 1$  alors  $a$  est répulsif et la divergence est au moins exponentielle.

*Démonstration.* Si  $|\varphi'(a)| < 1$  alors il existe  $k$  tel que  $|\varphi'(a)| < k < 1$  et par continuité il existe un voisinage  $V$  de  $a$  dans lequel  $|\varphi'(x)| < k$  pour tout  $x \in V$ . En utilisant le théorème des accroissements finis<sup>16</sup> nous avons

$$|x_n - a| = |\varphi(x_n) - a| \leq k|x_{n-1} - a| \quad (17.124)$$

et par récurrence

$$|x_n - a| \leq k^n |x_0 - a|. \quad (17.125)$$

Pour la seconde partie, nous supposons que  $|f'(a)| > 1$ . Il existe un voisinage  $V$  sur lequel  $f' > 1$ . La formule des accroissements finis donne une application  $\alpha: \mathbb{R} \rightarrow \mathbb{R}$  telle que

$$f(a + h) = f(a) + hf'(a) + \alpha(h) \quad (17.126)$$

et  $\alpha(h)/h \rightarrow 0$ . Nous restreignons  $V$  pour que tout  $x \in V$  nous ayons  $|\alpha(x - a)/(x - a)| + |f'(a)| < 1$ . Bref. Supposons que  $x_n \in V$  avec  $x_n \neq a$ . Nous avons alors

$$x_{n+1} = f(x_n) \quad (17.127a)$$

$$= f(a + (x_n - a)) \quad (17.127b)$$

$$= f(a) + (x_n - a)f'(a) + \alpha(x_n - a). \quad (17.127c)$$

Nous avons donc

$$\frac{|x_{n+1} - f(a)|}{|x_n - a|} \leq |f'(a)| + \frac{|\alpha(x_n - a)|}{|x_n - a|}. \quad (17.128)$$

Avec toutes les hypothèses que nous avons prises, nous avons

$$\frac{|x_{n+1} - f(a)|}{|x_n - a|} \leq 1, \quad (17.129)$$

et donc  $|x_{n+1} - f(a)| < |x_n - a|$ . □

**Remarque 17.32.**

Dans le cas  $|\varphi'(a)| = 1$ , nous ne pouvons rien conclure. Si  $\varphi(x) = \sin(x)$  nous avons  $\sin(x) < x$  et le point  $a = 0$  est attractif. À contrario, si  $\varphi(x) = \sinh(x)$  nous avons  $|\sinh(x)| > |x|$  et le point  $a = 0$  est répulsif.

**17.6.2 Picard**

DEFooRSLCooAsWisu

**Définition 17.33.**

Une application  $f: (X, d_X) \rightarrow (Y, d_Y)$  entre deux espaces métriques est une **contraction** si elle est  **$k$ -Lipschitz** pour un certain  $0 \leq k < 1$ , c'est-à-dire si pour tout  $x, y \in X$  nous avons

$$d_Y(f(x), f(y)) \leq kd_X(x, y). \quad (17.130)$$

16. Proposition 10.280.

ThoEPVCL

**Théorème 17.34** (Picard [563, 564]<sup>17</sup>).

Soit  $X$  un espace métrique complet et  $f: X \rightarrow X$  une application contractante, de constante de Lipschitz  $k$ . Alors  $f$  admet un unique point fixe, nommé  $\xi$ . Ce dernier est donné par la limite de la suite définie par récurrence

$$\begin{cases} x_0 \in X \\ x_{n+1} = f(x_n). \end{cases} \quad \begin{matrix} (17.131a) \\ (17.131b) \end{matrix}$$

De plus nous pouvons majorer l'erreur par

$$\|x_n - x\| \leq \frac{k^n}{1-k} \|x_n - x_{n-1}\| \leq \frac{k^n}{1-k} \|x_1 - x_0\|. \quad (17.132) \quad \text{EaKErdim}$$

Soit  $r > 0$ ,  $a \in X$  tels que la fonction  $f$  laisse la boule  $K = \overline{B(a, r)}$  invariante (c'est-à-dire que  $f$  se restreint à  $f: K \rightarrow K$ ). Nous considérons les suites  $(u_n)$  et  $(v_n)$  définies par

$$\begin{cases} u_0 = v_0 \in K \\ u_{n+1} = f(v_n), v_{n+1} \in B(u_n, \epsilon). \end{cases} \quad \begin{matrix} (17.133a) \\ (17.133b) \end{matrix}$$

Alors le point fixe  $\xi$  de  $f$  est dans  $K$  et la suite  $(v_n)$  satisfait l'estimation

$$\|v_n - \xi\| \leq \frac{k^n}{1-k} \|u_1 - u_0\| + \frac{\epsilon}{1-k}. \quad (17.134)$$

La première inégalité (17.132) donne une estimation de l'erreur calculable en cours de processus ; la seconde donne une estimation de l'erreur calculable avant de commencer.

*Démonstration.* Nous commençons par l'unicité du point fixe. Si  $a$  et  $b$  sont des points fixes, alors  $f(a) = a$  et  $f(b) = b$ . Par conséquent

$$\|f(a) - f(b)\| = \|a - b\|, \quad (17.135)$$

ce qui contredit le fait que  $f$  soit une contraction.

En ce qui concerne l'existence, notons que si la suite des  $x_n$  converge dans  $X$ , alors la limite est un point fixe. En effet en prenant la limite des deux côtés de l'équation  $x_{n+1} = f(x_n)$ , nous obtenons  $\xi = f(\xi)$ , c'est-à-dire que  $\xi$  est un point fixe de  $f$ . Notons que nous avons utilisé ici la continuité de  $f$ , laquelle est une conséquence du fait qu'elle soit Lipschitz. Nous allons donc porter nos efforts à prouver que la suite est de Cauchy (et donc convergente parce que  $X$  est complet). Nous commençons par prouver que  $\|x_{n+1} - x_n\| \leq k^n \|x_0 - x_1\|$ . En effet pour tout  $n$  nous avons

$$\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leq k \|x_n - x_{n-1}\|. \quad (17.136)$$

La relation cherchée s'obtient alors par récurrence. Soient  $q > p$ . En utilisant une somme télescopique,

$$\|x_q - x_p\| \leq \sum_{l=p}^{q-1} \|x_{l+1} - x_l\| \quad (17.137a)$$

$$\leq \left( \sum_{l=p}^{q-1} k^l \right) \|x_1 - x_0\| \quad (17.137b)$$

$$\leq \left( \sum_{l=p}^{\infty} k^l \right) \|x_1 - x_0\|. \quad (17.137c)$$

17. Il me semble qu'à la page 100 de [564], l'hypothèse H1 qui est prouvée ne prouve pas Hn dans le cas  $n = 1$ . Merci de m'écrire si vous pouvez confirmer ou infirmer. La preuve donnée ici ne contient pas cette « erreur ».

Étant donné que  $k < 1$ , la parenthèse est la queue d'une série qui converge, et donc tend vers zéro lorsque  $p$  tend vers l'infini.

En ce qui concerne les inégalités (17.132), nous refaisons une somme télescopique :

$$\|x_{n+p} - x_n\| \leq \|x_{n+p} - x_{n+p-1}\| + \cdots + \|x_{n+1} - x_n\| \quad (17.138a)$$

$$\leq k^p \|x_n - x_{n-1}\| + k^{p-1} \|x_n - x_{n-1}\| + \cdots + k \|x_n - x_{n-1}\| = k(1 + \cdots + k^{p-1}) \|x_n - x_{n-1}\| \quad (17.138b)$$

$$\leq \frac{k}{1-k} \|x_n - x_{n-1}\|. \quad (17.138c)$$

En prenant la limite  $p \rightarrow \infty$  nous trouvons

$$\|\xi - x_n\| \leq \frac{k}{1-k} \|x_n - x_{n-1}\| \leq \frac{k}{1-k} \|x_1 - x_0\|. \quad (17.139) \quad \text{EqUMVGW}$$

Nous passons maintenant à la seconde partie du théorème en supposant que  $f$  se restreigne en une fonction  $f: K \rightarrow K$ . D'abord  $K$  est encore un espace métrique complet, donc la première partie du théorème s'y applique et  $f$  y a un unique point fixe.

Nous allons montrer la relation par récurrence. Tout d'abord pour  $n = 1$  nous avons

$$\|v_1 - \xi\| \leq \|v_1 - u_1\| + \|u_1 - \xi\| \leq \epsilon + \frac{k}{1-k} \|u_1 - u_0\| \quad (17.140)$$

où nous avons utilisé l'estimation (17.139), qui reste valable en remplaçant  $x_1$  par  $u_1$ <sup>18</sup>. Nous pouvons maintenant faire la récurrence :

$$\|v_{n+1} - \xi\| \leq \|v_{n+1} - u_{n+1}\| + \|u_{n+1} - \xi\| \quad (17.141a)$$

$$\leq \epsilon + k \|v_n - \xi\| \quad (17.141b)$$

$$\leq \epsilon + k \left( \frac{k^n}{1-k} \|u_1 - u_0\| + \frac{\epsilon}{1-k} \right) = \frac{\epsilon}{1-k} + \frac{k^{n+1}}{1-k} \|u_1 - u_0\|. \quad (17.141c)$$

□

### Remarque 17.35.

Ce théorème comporte deux parties d'intérêts différents. La première partie est un théorème de point fixe usuel, qui sera utilisé pour prouver l'existence de certaines équations différentielles.

La seconde partie est intéressante d'un point de vue numérique. En effet, ce qu'elle nous enseigne est que si à chaque pas de calcul de la récurrence  $x_{n+1} = f(x_n)$  nous commettons une erreur d'ordre de grandeur  $\epsilon$ , alors le procédé (la suite  $(v_n)$ ) ne converge plus spécialement vers le point fixe, mais tend vers le point fixe avec une erreur majorée par  $\epsilon/(1-k)$ .

### Remarque 17.36.

Au final l'erreur minimale qu'on peut atteindre est de l'ordre de  $\epsilon$ . Évidemment si on commet une faute de calcul de l'ordre de  $\epsilon$  à chaque pas, on ne peut pas espérer mieux.

remIOHJm

### Remarque 17.37.

Si  $f$  elle-même n'est pas contractante, mais si  $f^p$  est contractante pour un certain  $p \in \mathbb{N}$  alors la conclusion du théorème de Picard reste valide et  $f$  a le même unique point fixe que  $f^p$ . En effet nommons  $x$  le point fixe de  $f : f^p(x) = x$ . Nous avons alors

$$f^p(f(x)) = f(f^p(x)) = f(x), \quad (17.142)$$

ce qui prouve que  $f(x)$  est un point fixe de  $f^p$ . Par unicité nous avons alors  $f(x) = x$ , c'est-à-dire que  $x$  est également un point fixe de  $f$ .

18. Elle n'est cependant pas spécialement valable si on remplace  $x_n$  par  $u_n$ .

ThoagJPZJ

**Théorème 17.38** (Équation de Fredholm).

Soit  $K: [a, b] \times [a, b] \rightarrow \mathbb{R}$  et  $\varphi: [a, b] \rightarrow \mathbb{R}$ , deux fonctions continues. Alors si  $\lambda$  est suffisamment petit, l'équation

$$f(x) = \lambda \int_a^b K(x, y)f(y)dy + \varphi(x) \quad (17.143)$$

admet une unique solution qui sera de plus continue sur  $[a, b]$ .

*Démonstration.* Nous considérons l'ensemble  $\mathcal{F}$  des fonctions continues  $[a, b] \rightarrow [a, b]$  muni de la norme uniforme. Le lemme 12.384 implique que  $\mathcal{F}$  est complet. Nous considérons l'application  $\Phi: \mathcal{F} \rightarrow \mathcal{F}$  donnée par

$$\Phi(f)(x) = \lambda \int_a^b K(x, y)f(y)dy + \varphi(x). \quad (17.144)$$

Nous montrons que  $\Phi^p$  est une application contractante pour un certain  $p$ . Pour tout  $x \in [a, b]$  nous avons

$$\|\Phi(f) - \Phi(g)\|_\infty \leq \|\Phi(f)(x) - \Phi(g)(x)\| \quad (17.145a)$$

$$= |\lambda| \left\| \int_a^b K(x, y)(f(y) - g(y))dy \right\| \quad (17.145b)$$

$$\leq |\lambda| \|K\|_\infty |b - a| \|f - g\|_\infty \quad (17.145c)$$

Nous choisissons  $\lambda$  assez petit pour avoir  $|\lambda| \|K\|_\infty |b - a| < 1$ . Dans ce cas, l'application  $\Phi$  est une contraction. Elle possède donc un unique point fixe par le théorème de Picard 17.34.  $\square$

## 17.7 Théorèmes de point fixes et équations différentielles

### 17.7.1 Théorème de Cauchy-Lipschitz

Nous démontrons ici deux théorèmes de Cauchy-Lipschitz. De nombreuses propriétés annexes seront démontrées dans le chapitre sur les équations différentielles, section 32.8.

Le théorème de Cauchy-Arzella 20.34 sera pour plus tard parce qu'il utilise Schauder 20.33.

ThokUUlgU

**Théorème 17.39** (Cauchy-Lipschitz[565, 566]).

Nous considérons l'équation différentielle <sup>XtiXON</sup>

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \quad (17.146a)$$

$$(17.146b)$$

avec  $f: U = I \times \Omega \rightarrow \mathbb{R}^n$  où  $I$  est ouvert dans  $\mathbb{R}$  et  $\Omega$  ouvert dans  $\mathbb{R}^n$ . Nous supposons que  $f$  est continue sur  $U$  et localement Lipschitz<sup>19</sup> par rapport à  $y$ .

Alors il existe un intervalle  $J \subset I$  sur lequel la solution au problème est unique. De plus toute solution du problème est une restriction de cette solution à une partie de  $J$ . La solution sur  $J$  (dite « solution maximale ») est de classe  $C^1$ .

*Démonstration.* Nous divisons la preuve en plusieurs étapes (même pas toutes simples).

#### (1) Cylindre de sécurité

Précisons l'espace fonctionnel  $\mathcal{F}$  adéquat. Soient  $V$  et  $W$  les voisinages de  $t_0$  et  $y_0$  sur lesquels  $f$  est localement Lipschitz. Nous considérons les quantités suivantes :

$$(1a) \quad M = \sup_{V \times W} f;$$

$$(1b) \quad r > 0 \text{ tel que } \overline{B(y_0, r)} \subset W$$

$$(1c) \quad T > 0 \text{ tel que } \overline{B(t_0, T)} \subset V \text{ et } T < r/M.$$

19. Définition 12.345. Notons que nous ne supposons pas que  $f$  soit une contraction.

Nous considérons alors l'ensemble

$$\mathcal{F} = C^0(\overline{B(t_0, T)}, \overline{B(y_0, r)}) \quad (17.147)$$

que nous munissons de la norme uniforme. Par le lemme 12.384 l'espace  $(\mathcal{F}, \|\cdot\|_\infty)$  est complet.

(2) **Une application  $\Phi: \mathcal{F} \rightarrow \mathcal{F}$**

Si  $y$  est une solution de l'équation différentielle considérée, elle vérifie<sup>20</sup>

$$y(t) = y_0 + \int_{t_0}^t f(u, y(u)) du. \quad (17.148) \quad \text{EqPGLwcl}$$

Ceci nous incite à considérer l'opérateur  $\Phi: \mathcal{F} \rightarrow \mathcal{F}$  défini par

$$\Phi(y)(t) = y_0 + \int_{t_0}^t f(u, y(u)) du. \quad (17.149)$$

Pour que l'application  $\Phi$  soit utile nous devons montrer que pour tout  $y \in \mathcal{F}$ ,

- l'application  $\Phi(y)$  est bien définie,
- pour tout  $t \in \overline{B(t_0, T)}$  nous avons  $\Phi(y)(t) \in \overline{B(y_0, r)}$ ,
- l'application  $\Phi(y): \overline{B(t_0, T)} \rightarrow \overline{B(y_0, r)}$  est continue.

Attention : nous ne prétendons pas que  $\Phi$  elle-même soit continue. C'est parti.

(2a)  **$\Phi(y)$  est bien définie**

Il faut montrer que l'intégrale converge. Le calcul de  $\Phi(y)(t)$  ne se fait qu'avec  $t \in \overline{B(t_0, T)}$ . Vu que  $u$  prend ses valeurs dans  $[t_0, t]$  et que  $y \in \mathcal{F}$ , le nombre  $y(u)$  est toujours dans  $\overline{B(y_0, r)}$ . Ceci pour dire que dans l'intégrale, la fonction  $f$  n'est considérée que sur  $[t_0, t] \times \overline{B(y_0, r)} \subset V \times W$ . La fonction  $f$  est donc uniformément majorable, et l'intégrale ne pose pas de problèmes.

(2b)  **$\Phi(y)(t) \in \overline{B(y_0, r)}$**

Prouvons que  $\Phi(y)(t) \in \overline{B(y_0, r)}$ . Pour cela, notons que

$$|\Phi(y)(t) - y_0| \leq \int_{t_0}^t |f(u, y(u))| du \leq |t - t_0| \|f\|_\infty. \quad (17.150)$$

Étant donné que  $t \in \overline{B(t_0, T)}$  nous avons  $|t - t_0| \leq r/M$  et donc  $|\Phi(y)(t) - y_0| \leq r$ .

(2c)  **$\Phi(y)$  est continue**

Nous pourrions invoquer le théorème 17.15, mais nous allons le faire à la main. Soit  $s_0 \in \overline{B(t_0, T)}$  et prouvons que  $\Phi(y)$  est continue en  $s_0$ . Pour cela nous prenons  $s \in \overline{B(s_0, \delta)}$  et nous calculons :

$$|\Phi(y)(s) - \Phi(y)(s_0)| \leq \int_{s_0}^s |f(u, y(u))| du \leq |s_0 - s| \|f\|_\infty. \quad (17.151)$$

C'est le fait que  $f$  soit bornée dans le cylindre de sécurité qui fait en sorte que cela tende vers zéro lorsque  $s \rightarrow s_0$ .

L'équation (17.148) signifie que  $y$  est un point fixe de  $\Phi$ . L'espace  $\mathcal{F}$  étant complet, le théorème de point fixe de Picard (théorème 17.34) s'applique. Nous allons montrer qu'il existe un  $p \in \mathbb{N}$  tel que  $\Phi^p$  soit contractante. Par conséquent  $\Phi^p$  aura un unique point fixe qui sera également unique point fixe de  $\Phi$  par la remarque 17.37.

(3) **Contractante**

Prouvons donc que  $\Phi^p$  est contractante pour un certain  $p$ . Pour cela nous commençons par montrer la formule suivante par récurrence :

$$\|\Phi^p(x)(t) - \Phi^p(y)(t)\| \leq \frac{k^p |t - t_0|^p}{p!} \|x - y\|_\infty \quad (17.152) \quad \text{EqBAdKxT}$$

20. C'est le théorème fondamental du calcul intégral 14.288.



pour tout  $x, y \in \mathcal{F}$ , et pour tout  $t \in \overline{B(t_0, T)}$ . Pour  $p = 0$  la formule (17.152) est vérifiée parce que  $\|x - y\|_\infty$  est le supremum de  $\|x(t) - y(t)\|$  pour  $t \in \overline{B(t_0, T)}$ . Supposons que la formule soit vraie pour  $p$  et calculons pour  $p + 1$ . Pour tout  $t \in \overline{B(t_0, T)}$  nous avons

$$\|\Phi^{p+1}(x)(t) - \Phi^{p+1}(y)(t)\| \leq \left| \int_{t_0}^t \|f(u, \Phi^p(x)(u)) - f(u, \Phi^p(y)(u))\| du \right| \quad (17.153a)$$

$$\leq \left| \int_{t_0}^t k \|\Phi^p(x)(u) - \Phi^p(y)(u)\| du \right| \quad (17.153b)$$

$$\leq \left| \int_{t_0}^t k \frac{k^p |t - t_0|^p}{p!} \|x - y\|_\infty du \right| \quad (17.153c)$$

$$= \frac{k^{p+1} |t - t_0|^{p+1}}{(p+1)!} \|x - y\|_\infty. \quad (17.153d)$$

Justifications :

- (17.153b) parce que  $f$  est Lipschitz.
- (17.153c) par hypothèse de récurrence.

La formule (17.152) est maintenant établie. Nous pouvons maintenant montrer que  $\Phi^p$  est une contraction pour un certain  $p$ . Pour tout  $t \in \overline{B(t_0, T)}$  nous avons

$$\|\Phi^p(x)(t) - \Phi^p(y)(t)\| \leq \frac{k^p}{p!} |t - t_0|^p \|x - y\|_\infty \leq \frac{k^p T^p}{p!} \|x - y\|_\infty \quad (17.154)$$

où nous avons utilisé le fait que  $|t - t_0|^p < T^p$ . En prenant le supremum sur  $t$  des deux côtés il vient

$$\|\Phi^p(x) - \Phi^p(y)\|_\infty \leq \frac{k^p T^p}{p!} \|x - y\|_\infty. \quad (17.155)$$

Le membre de droite tend vers zéro lorsque  $p \rightarrow \infty$  parce que  $k^p T^p / p! \rightarrow 0$ <sup>21</sup>. Nous concluons donc que  $\Phi^p$  est une contraction pour un certain  $p$ .

#### (4) Conclusion

L'unique point fixe de  $\Phi$  est alors l'unique solution continue de l'équation différentielle (17.146). Par ailleurs l'équation elle-même  $y' = f(t, y)$  demande implicitement que  $y$  soit dérivable et donc continue. Nous concluons que l'unique point fixe de  $\Phi$  est l'unique solution de l'équation différentielle donnée. Cette dernière est automatiquement  $C^1$  parce que si  $y$  est continue alors  $u \mapsto f(u, y(u))$  est continue, c'est-à-dire que  $y'$  est continue.

#### (5) Unicité

Nous passons maintenant à la partie « prolongement maximum » du théorème. Soient  $x_1$  et  $x_2$  deux solutions maximales du problème (17.146) sur des intervalles  $I_1$  et  $I_2$  respectivement. Les intervalles  $I_1$  et  $I_2$  contiennent  $\overline{B(t_0, T)}$  sur lequel  $x_1 = x_2$  par unicité.

Nous allons maintenant montrer que pour tout  $t \geq t_0$  pour lequel  $x_1$  ou  $x_2$  est défini,  $x_1(t)$  et  $x_2(t)$  sont définis et sont égaux. Le raisonnement sur  $t \leq t_0$  est similaire.

Supposons que l'ensemble des  $t \geq t_0$  tels que  $x_1 = x_2$  soit ouvert à droite, c'est-à-dire soit de la forme  $[t_0, b[$ . Dans ce cas, soit  $x_1$ , soit  $x_2$  (soit les deux) cesse d'exister en  $b$ . En effet si nous avions les fonctions  $x_i$  sur  $[t_0, b + \epsilon[$  alors l'équation  $x_1 = x_2$  définirait un fermé dans  $[t_0, b + \epsilon[$ . Supposons pour fixer les idées que  $x_1$  cesse d'exister : le domaine de  $x_1$  (parmi les  $t \geq 0$ ) est  $[t_0, b[$  et sur ce domaine nous avons  $x_1 = x_2$ . Dans ce cas  $x_1$  pourrait être prolongé en  $x_2$  au-delà de  $b$ . Si  $x_1$  et  $x_2$  s'arrêtent d'exister en même temps en  $b$ , alors nous avons bien  $x_1 = x_2$ .

Nous devons donc traiter le cas où  $x_1 = x_2$  sur  $[t_0, b]$  alors que  $x_1$  et  $x_2$  existent sur  $[t_0, b + \epsilon[$  pour un certain  $\epsilon$ .

21. Parce que les factorielles vont plus vite que les puissances.

Nous pouvons appliquer le théorème d'existence locale au problème

$$\begin{cases} y' = f(t, y) \\ y(b) = x_1(b). \end{cases} \quad (17.156a)$$

$$\begin{cases} y(b) = x_2(b). \end{cases} \quad (17.156b)$$

Il existe un voisinage de  $b$  sur lequel la solution est unique. Sur ce voisinage nous devons donc avoir  $x_1 = x_2$ , ce qui contredit le fait que  $x_1 \neq x_2$  en dehors de  $[t_0, b]$ .

Donc  $x_1$  et  $x_2$  existent et sont égaux sur, au moins  $I_1 \cup I_2$ .

□

Le théorème de Cauchy-Lipschitz donne existence et unicité d'une solution maximale. Cependant cette solution peut ne pas exister partout où les hypothèses sur  $f$  sont remplies. En d'autres termes, il peut arriver que  $f$  soit Lipschitz jusqu'à  $t_1$ , mais que la solution maximale ne soit définie que jusqu'en  $t_2 < t_1$ . Ce cas fait l'objet du théorème d'explosion en temps fini 32.23.

Sous quelques hypothèses, nous pouvons nous assurer de l'existence d'une solution unique sur tout  $\mathbb{R}$ .

THOooZIVRooPSWMxg

**Théorème 17.40** (Cauchy-Lipschitz global[567, 126]).

Soit un intervalle  $I$  de  $\mathbb{R}$ ,  $y_0 \in \mathbb{R}^n$ ,  $t_0 \in I$  et une fonction continue  $f: I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  telle que pour tout compact  $K$  dans  $I$ , il existe  $k > 0$  tel que

$$\|f(t, y_1) - f(t, y_2)\| \leq k\|y_1 - y_2\| \quad (17.157)$$

pour tout  $t \in K$  et  $y_1, y_2 \in \mathbb{R}^n$ .

Alors le problème |EQSooBNREooUTfbMH

$$\begin{cases} y'(t) = f(t, y(t)) \\ y(t_0) = y_0 \end{cases} \quad (17.158a)$$

$$\begin{cases} y(t_0) = y_0 \end{cases} \quad (17.158b)$$

possède une unique solution  $y: I \rightarrow \mathbb{R}^n$  sur  $I$ .

*Démonstration.* Soit un intervalle compact  $K$  dans  $I$  et contenant  $t_0$ . Nous notons  $\ell$  le diamètre de  $K$ . Sur l'espace  $E = C^0(K, \mathbb{R}^n)$  nous considérons la topologie uniforme :  $(E, \|\cdot\|_\infty)$ . C'est un espace complet par le lemme 12.384 (nous utilisons le fait que  $\mathbb{R}^n$  soit complet, proposition 1.508). Nous allons utiliser l'application suivante :

$$\begin{aligned} \Phi: E &\rightarrow E \\ \Phi(y)(t) &= y_0 + \int_{t_0}^t f(s, y(s)) ds \end{aligned} \quad (17.159)$$

EQooJUTBooILBKooE

Démontrons quelques faits à propos de  $\Phi$ .

- (1)  **$\Phi$  est bien définie** Nous devons commencer par prouver que cette application est bien définie. Si  $y \in E$  alors  $f$  et  $y$  sont continues ; l'application  $s \mapsto f(s, y(s))$  est donc également continue. L'intégrale de cette fonction sur le compact  $[t_0, t]$  ne pose alors pas de problèmes. En ce qui concerne la continuité de  $\Phi(y)$  sous l'hypothèse que  $y$  soit continue,

$$\|\Phi(y)(t) - \Phi(y)(t')\| \leq \int_t^{t'} \|f(s, y(s))\| ds \leq M|t - t'| \quad (17.160)$$

où  $M$  est une majoration de  $\|s \mapsto f(s, y(s))\|_{\infty, K}$ .

- (2) **Si  $y$  est solution alors  $\Phi(y) = y$**

Supposons que  $y$  soit une solution de l'équation différentielle (17.158). Alors, vu que  $y'(t) = f(t, y(t))$  nous avons :

$$y(t) = y_0 + \int_{t_0}^t y'(s) ds = y_0 + \int_{t_0}^t f(s, y(s)) ds = \Phi(y)(t). \quad (17.161)$$

(3) **Si  $\Phi(y) = y$  alors  $y$  est solution**

Nous avons, pour tout  $t$  :

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) ds. \quad (17.162)$$

Le membre de droite est dérivable par rapport à  $t$ , et la dérivée est  $f(t, y(t))$ . Donc le membre de gauche est également dérivable et nous avons bien

$$y'(t) = f(t, y(t)). \quad (17.163)$$

De plus  $y(t_0) = y_0 + \int_{t_0}^{t_0} \dots = y_0$ .

Nous sommes encore avec  $K$  compact et  $E = C^0(K, \mathbb{R}^n)$  muni de la norme uniforme. Nous allons montrer que  $\Phi$  est une contraction de  $E$  pour une norme bien choisie.

(1) **Une norme sur  $E$**  Pour  $y \in E$  nous posons

$$\|y\|_k = \max_{t \in K} (e^{-k|t-t_0|} \|y(t)\|). \quad (17.164)$$

Ce maximum est bien défini et fini, parce que dedans, la fonction de  $t$  est une fonction continue sur le compact  $K$ . C'est également une norme parce que si  $\|y\|_k = 0$  alors  $e^{-k|t-t_0|} \|y(t)\| = 0$  pour tout  $t$ . Étant donné que l'exponentielle ne s'annule pas,  $\|y(t)\| = 0$  pour tout  $t$ .

(2) **Équivalence de norme** Nous montrons que les normes  $\|\cdot\|_k$  et  $\|\cdot\|_\infty$  sont équivalentes<sup>22</sup> :

$$\|y\|_\infty e^{-k\ell} \leq \|y\|_k \leq \|y\|_\infty \quad (17.165)$$

pour tout  $y \in E$ . Pour la première inégalité,  $\ell \geq |t - t_0|$  pour tout  $t \in K$ , et  $k > 0$ , donc

$$\|y(t)\| e^{-k\ell} \leq e^{-k|t-t_0|} \|y(t)\|. \quad (17.166)$$

En prenant le maximum des deux côtés,  $\|y\|_\infty e^{-k\ell} \leq \|y\|_k$ .

En ce qui concerne la seconde inégalité dans (17.165),  $k|t - t_0| \geq 0$  et donc  $e^{-k|t-t_0|} < 1$ .

Puisque les normes  $\|\cdot\|_\infty$  et  $\|\cdot\|_k$  sont équivalentes, l'espace  $(E, \|\cdot\|_k)$  est tout autant complet que  $(E, \|\cdot\|_\infty)$ . Nous démontrons à présent que  $\Phi$  est une contraction dans  $(E, \|\cdot\|_k)$ .

Soient  $y, z \in E$ . Si  $t \geq t_0$  nous avons

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leq \int_{t_0}^t \|f(s, y(s)) - f(s, z(s))\| ds \quad (17.167a)$$

$$\leq k \int_{t_0}^t \|y(s) - z(s)\| ds. \quad (17.167b)$$

Il convient maintenant de remarquer que

$$\|y(t)\| = e^{-k|t-t_0|} e^{k|t-t_0|} \|y(t)\| \leq \|y\|_k e^{k|t-t_0|}. \quad (17.168)$$

Nous pouvons avec ça prolonger les inégalités (17.167) par

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leq k \|y - z\|_k \int_{t_0}^t e^{k|s-t_0|} ds = k \|y - z\|_k \int_{t_0}^t e^{k(s-t_0)} ds \quad (17.169)$$

où nous avons utilisé notre supposition  $t \geq t_0$  pour éliminer les valeurs absolues. L'intégrale peut être calculée explicitement, mais nous en sommes arrivés à un niveau de fainéantise tellement inconcevable que

---

22. Définition 10.54

```

1 SageMath version 7.3, Release Date: 2016-08-04
2 Type "notebook()" for the browser-based notebook interface.
3 Type "help()" for help.
4
5
6 sage: var('a,b,k')
7 (a, b, k)
8 sage: f(x)=exp(-k*x)
9 sage: f.integrate(x,a,b)
10 e^(-a*k)/k - e^(-b*k)/k

```

tex/sage/sageSnip014.sage

Au final, si  $t \geq t_0$ ,

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leq \|y - z\|_k (e^{k(t-t_0)} - 1). \quad (17.170)$$

Si  $t \leq t_0$ , il faut retourner les bornes de l'intégrale avant d'y faire rentrer la norme parce que  $\|\int_0^1 f\| \leq \int_0^1 \|f\|$ , mais ça ne marche pas avec  $\|\int_1^0 f\|$ . Pour  $t \leq t_0$  tout le calcul donne

$$\|\Phi(y)(t) - \Phi(z)(t)\| \leq \|y - z\|_k (e^{k(t_0-t)} - 1). \quad (17.171)$$

Les deux inéquations sont valables a fortiori en mettant des valeurs absolues dans l'exponentielle, de telle sorte que pour tout  $t \in K$  nous avons

$$e^{-k|t_0-t|} \|\phi(y)(t) - \phi(z)(t)\| \leq \|y - z\|_k (1 - e^{-k|t_0-t|}). \quad (17.172)$$

En prenant le supremum sur  $t$ ,

$$\|\Phi(y) - \Phi(z)\|_k \leq \|y - z\|_k (1 - e^{-k\ell}), \quad (17.173)$$

mais  $0 < (1 - e^{-k\ell}) < 1$ , donc  $\Phi$  est contractante pour la norme  $\|\cdot\|_k$ . Comme  $(E, \|\cdot\|_k)$  est complet, l'application  $\Phi$  y a un unique point fixe par le théorème de Picard 17.34.

Ce point fixe est donc l'unique solution de l'équation différentielle de départ.

- (1) **Existence et unicité sur  $I$**  Il nous reste à prouver que la solution que nous avons trouvée existe sur  $I$  : jusqu'à présent nous avons démontré l'existence et l'unicité sur n'importe quel compact dans  $I$ .

Soit une suite croissante de compacts  $K_n$  contenant  $t_0$  (par exemple une suite exhaustive comme celle du lemme 7.358). Nous avons en particulier

$$I = \bigcup_{n=0}^{\infty} K_n. \quad (17.174)$$

- (2) **Existence sur  $I$**  Soit  $y_n$  l'unique solution sur  $K_n$ . Il suffit de poser

$$y(t) = y_n(t) \quad (17.175)$$

pour  $n$  tel que  $t \in K_n$ . Cette définition fonctionne parce que si  $t \in K_n \cap K_m$ , il y a forcément un des deux qui est inclus dans l'autre et le résultat d'unicité sur le plus grand des deux donne  $y_n(t) = y_m(t)$ .

- (3) **Unicité sur  $I$**

Soient  $y$  et  $z$  des solutions sur  $I$  ; puisque  $I$  n'est pas spécialement compact, le travail fait plus haut ne permet pas de conclure que  $y = z$ .

Soit  $t \in I$ . Alors  $t \in K_n$  pour un certain  $n$  et  $y$  et  $z$  sont des solutions sur  $K_n$  qui est compact. L'unicité sur  $K_n$  donne  $y(t) = z(t)$ .

□

**17.41.**

Il y a d'autres moyens de prouver qu'une solution existe globalement sur  $\mathbb{R}$ . Si  $f$  est globalement bornée, le théorème d'explosion en temps fini donne quelques garanties, voir 32.25.

Le théorème suivant donne une version du théorème de Cauchy-Lipschitz lorsque la fonction  $f$  dépend d'un paramètre. Ce théorème n'utilise rien de fondamentalement nouveau. Nous le donnons seulement pour montrer que l'on peut choisir l'espace  $\mathcal{F}$  de façon un peu maligne pour élargir le résultat. Si vous voulez un théorème de Cauchy-Lipschitz avec paramètre vraiment intéressant, allez voir le théorème 32.40.

THOooDTCWooSPKeYu

**Théorème 17.42** (Cauchy-Lipschitz avec paramètre[1, 568]).

Soit un intervalle ouvert  $I$  de  $\mathbb{R}$ , un connexe ouvert  $\Omega$  de  $\mathbb{R}^n$  et un intervalle ouvert  $\Lambda$  de  $\mathbb{R}^d$ . Soit une fonction  $f: I \times \Omega \times \Lambda \rightarrow \mathbb{R}^n$  continue et localement Lipschitz en  $\Omega$ . Soient  $t_0 \in I$ ,  $y_0 \in \Omega$  et  $\lambda_0 \in \Lambda$ . Il existe un voisinage compact de  $(t_0, y_0, \lambda_0)$  sur lequel le problème

$$\begin{cases} y'_\lambda(t) = f(t, y_\lambda(t), \lambda) \\ y_\lambda(t_0) = y_0 \end{cases} \quad (17.176a)$$

$$(17.176b)$$

possède une unique solution. De plus  $(t, \lambda) \mapsto y_\lambda(t)$  est continue<sup>23</sup>.

*Idée rapide de la preuve.*

**⚠ Avertissement/question au lecteur !! 17.43**

*Ceci est une idée de la preuve. Je n'ai pas vérifié toutes les étapes. Soyez prudent.*

D'abord nous avons un voisinage compact  $V \times \overline{B(y_0, r)} \times \Lambda_0$  de  $(t_0, y_0, \lambda_0)$  sur lequel  $f$  est bornée. Ensuite nous récrivons l'équation différentielle sous la forme

$$\begin{cases} \frac{\partial y}{\partial t}(t, \lambda) = f(t, y(t, \lambda), \lambda) \\ y(t_0, \lambda) = y_0. \end{cases} \quad (17.177a)$$

$$(17.177b)$$

pour une fonction  $y: V \times \Lambda_0 \rightarrow \mathbb{R}^n$ .

Nous posons  $\mathcal{F} = C^0(V \times \Lambda_0, \mathbb{R}^n)$  et nous y définissons l'application

$$\begin{aligned} \Phi: \mathcal{F} &\rightarrow \mathcal{F} \\ \Phi(y)(t, \lambda) &= y_0 + \int_{t_0}^t f(s, y(s, \lambda), \lambda) ds. \end{aligned} \quad (17.178)$$

Il y a plein de vérifications à faire[568], mais je parie que  $\Phi$  est bien définie, et qu'une de ses puissances est une contraction de  $(\mathcal{F}, \|\cdot\|_\infty)$ . L'unique point fixe est une solution de notre problème et est dans  $C^0$ , donc  $(t, \lambda) \mapsto y(t, \lambda) = y_\lambda(t)$  est de classe  $C^0$ , c'est-à-dire continue.  $\square$

**17.44.**

Ce théorème marque un peu la limite de ce que l'on peut faire avec la méthode des points fixes dans le cadre de Cauchy-Lipschitz : nous sommes limités à la continuité de la solution parce que les espaces  $C^p$  ne sont pas complets<sup>24</sup>. Il n'y a donc pas d'espoir d'adapter la méthode pour prouver que si  $f$  est de classe  $C^p$  alors  $(t, \lambda) \mapsto y_\lambda(t)$  est de classe  $C^p$ . On peut, à  $\lambda$  fixé, prouver que  $t \mapsto y_\lambda(t)$  est de classe  $C^p$  (utiliser une récurrence), mais pas plus.

La régularité  $C^1$  de  $y$  par rapport à la condition initiale sera l'objet du théorème 32.36. Ce résultat n'est vraiment pas facile et utilise des ingrédients bien autres qu'un point fixe. Ensuite la régularité  $C^p$  par rapport à la condition initiale et par rapport à un paramètre seront presque des cadeaux (proposition 32.38 et 32.40).

23. Ici, la surprise est que ce soit continu par rapport à  $\lambda$ . Le fait qu'elle le soit par rapport à  $t$  est clair depuis le départ, parce que ce n'est finalement rien d'autre que le Cauchy-Lipschitz vieux et connu.

24. Par exemple, le théorème de Stone-Weierstrass 12.446 nous dit que la limite uniforme de polynômes (de classe  $C^\infty$ ) peut n'être que continue. Voir aussi le thème 37.

**Exemple 17.45** ([569]).

Nous savons que le théorème de Picard permet de trouver le point fixe par itération de la contraction à partir d'un point quelconque. Tentons donc de résoudre

$$\begin{cases} y'(t) = y(t) \\ y(0) = 1 \end{cases} \quad (17.179a)$$

$$(17.179b)$$

dont nous savons depuis l'enfance que la solution est l'exponentielle<sup>25</sup>. Partons donc de la fonction constante  $y_0 = 1$ , et appliquons la contraction (17.159) :

$$u_1 = 1 + \int_0^1 u_0(s) ds = 1 + t. \quad (17.180)$$

Ensuite

$$u_2 = 1 + \int_0^t (1 + s) ds = 1 + t + \frac{t^2}{2}. \quad (17.181)$$

Et on voit que les itérations suivantes vont donner l'exponentielle.

Nous sommes évidemment en droit de se dire que nous avons choisi un bon point de départ. Tentons le coup avec une fonction qui n'a rien à voir avec l'exponentielle :  $u_0(x) = \sin(x)$ .

Le programme suivant permet de faire de belles investigations numériques en partant d'à peu près n'importe quelle fonction :

```

1  #! /usr/bin/sage -python
2  # -*- coding: utf8 -*-
3
4  from sage.all import *
5
6  x=var('x')
7
8  def Phi(f):
9      prim=f.integrate()
10     return 1+prim(x)-prim(0)
11
12 f=sin(x)
13
14 for i in range(1,30):
15     print(i,f)
16     f=Phi(f)
17
18 g=f(x)-exp(x)
19 plot(g,(x,-10,10)).show()
```

tex/sage/picard\_exp.py

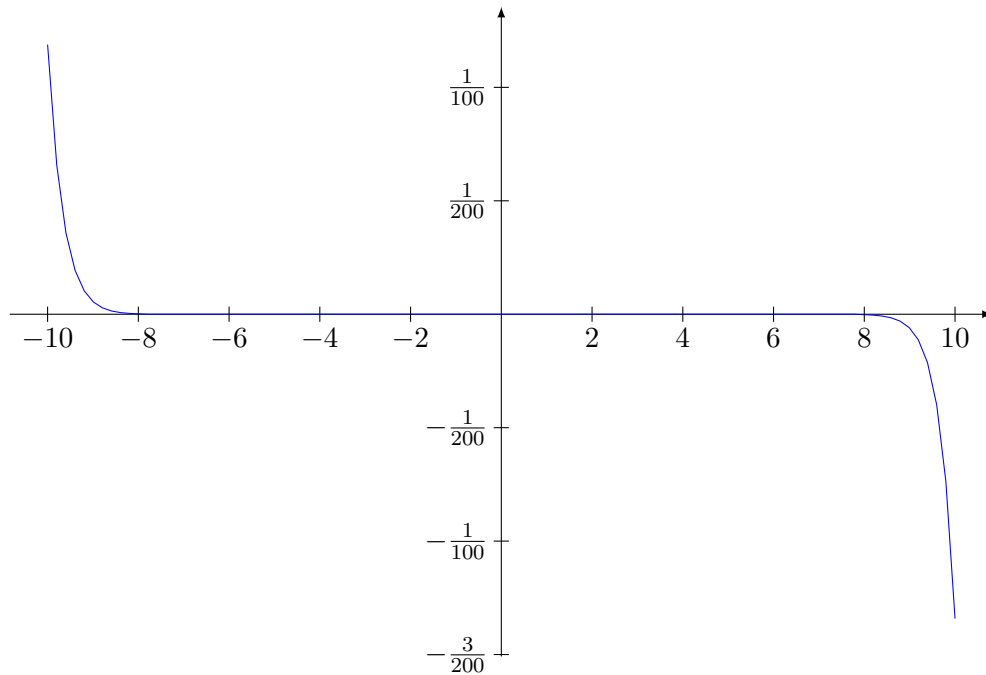
Ce programme fait 30 itérations depuis la fonction  $\sin(x)$  pour tenter d'approximer  $\exp(x)$ . Pour donner une idée, après 7 itérations nous avons la fonction suivante :

$$\frac{1}{60}x^5 + \frac{1}{24}x^4 + \frac{1}{2}x^2 + 2x - \sin(x) + 1. \quad (17.182)$$

Nous voyons que les coefficients sont des factorielles, mais pas toujours celles correspondantes à la puissance, et qu'il manque certains termes par rapport au développement de l'exponentielle que nous connaissons. Bref, le polynôme qui se met en face de  $\sin(x)$  s'adapte tout seul pour compenser.

Et après 30 itérations, ça donne quoi ? Voici un graphe de l'erreur entre  $u_{30}(x)$  et  $\exp(30)$  :

25. Voir par exemple le théorème 15.84.



Pour donner une idée,  $\exp(10) \simeq 22000$ . Donc il y a une faute de 0.01 sur 22000. Pas mal.

△

## 17.8 Théorèmes d'inversion locale et de la fonction implicite

### 17.8.1 Mise en situation

Dans un certain nombre de situation, il n'est pas possible de trouver des solutions explicites aux équations qui apparaissent. Néanmoins, l'existence « théorique » d'une telle solution est souvent déjà suffisante. C'est l'objet du théorème de la fonction implicite.

Prenons par exemple la fonction sur  $\mathbb{R}^2$  donnée par

$$F(x, y) = x^2 + y^2 - 1. \quad (17.183)$$

Nous pouvons bien entendu regarder l'ensemble des points donnés par  $F(x, y) = 0$ . C'est le cercle dessiné à la figure 17.1.

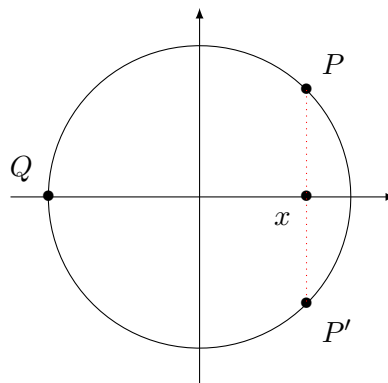


FIGURE 17.1: Un cercle pour montrer l'intérêt de la fonction implicite. Si on donne  $x$ , nous ne pouvons pas savoir si nous parlons de  $P$  ou de  $P'$ .

LabelFigCercleImplicite

Nous ne pouvons pas donner le cercle sous la forme  $y = y(x)$  à cause du  $\pm$  qui arrive quand on prend la racine carrée. Mais si on se donne le point  $P$ , nous pouvons dire que *autour de  $P$* , le cercle est la fonction

$$y(x) = \sqrt{1 - x^2}. \quad (17.184)$$

Tandis que autour du point  $P'$ , le cercle est la fonction

$$y(x) = -\sqrt{1-x^2}. \quad (17.185)$$

Autour de ces deux points, donc, le cercle est donné par une fonction. Il n'est par contre pas possible de donner le cercle autour du point  $Q$  sous la forme d'une fonction.

Ce que nous voulons faire, en général, est de voir si l'ensemble des points tels que

$$F(x_1, \dots, x_n, y) = 0 \quad (17.186)$$

peut être donné par une fonction  $y = y(x_1, \dots, x_n)$ . En d'autres termes, est-ce qu'il existe une fonction  $y(x_1, \dots, x_n)$  telle que

$$F(x_1, \dots, x_n, y(x_1, \dots, x_n)) = 0. \quad (17.187)$$

Plus généralement, soit une fonction

$$\begin{aligned} F: D \subset \mathbb{R}^n \times \mathbb{R}^m &\rightarrow \mathbb{R}^m \\ (x, y) &\mapsto (F_1(x, y), \dots, F_m(x, y)) \end{aligned} \quad (17.188)$$

avec  $x = (x_1, \dots, x_n)$  et  $y = (y_1, \dots, y_m)$ . Pour chaque  $x$  fixé, on s'intéresse aux solutions du système de  $m$  équations  $F(x, y) = 0$  pour les inconnues  $y$ ; en particulier, on voudrait pouvoir écrire  $y = \varphi(x)$  vérifiant  $F(x, \varphi(x)) = 0$ .

### 17.8.2 Théorème d'inversion locale

LemGZoqknC

**Lemme 17.46** ([404]).

Soit  $E$  un espace de Banach (métrique complet) et  $\mathcal{O}$  un ouvert de  $E$ . Nous considérons une  $\lambda$ -contraction  $\varphi: \mathcal{O} \rightarrow E$ . Alors l'application

$$f: x \mapsto x + \varphi(x) \quad (17.189)$$

est un homéomorphisme entre  $\mathcal{O}$  et un ouvert de  $E$ . De plus  $f^{-1}$  est Lipschitz de constante plus petite ou égale à  $(1 - \lambda)^{-1}$ .

Cette proposition utilise le théorème de point fixe de Picard 17.34, et sera utilisée pour démontrer le théorème d'inversion locale 17.47.

*Démonstration.* Soient  $x_1, x_2 \in \mathcal{O}$ . Nous posons  $y_1 = f(x_1)$  et  $y_2 = f(x_2)$ . En vertu de l'inégalité de la proposition 7.213 nous avons subEqEBjsBfz

$$\|f(x_2) - f(x_1)\| = \|x_2 + \varphi(x_2) - x_1 - \varphi(x_1)\| \quad (17.190a)$$

$$\geq \|x_2 - x_1\| - \|\varphi(x_2) - \varphi(x_1)\| \quad (17.190b)$$

$$\geq (1 - \lambda)\|x_2 - x_1\|. \quad (17.190c)$$

À la dernière ligne les valeurs absolues sont enlevées parce que nous savons que ce qui est à l'intérieur est positif. Cela nous dit d'abord que  $f$  est injective parce que  $f(x_2) = f(x_1)$  implique  $x_2 = x_1$ . Donc  $f$  est inversible sur son image. Nous posons  $A = f(\mathcal{O})$  et nous devons prouver que que  $f^{-1}: A \rightarrow \mathcal{O}$  est continue, Lipschitz de constante majorée par  $(1 - \lambda)^{-1}$  et que  $A$  est ouvert.

Les inéquations (17.190) nous disent que

$$\|f^{-1}(y_1) - f^{-1}(y_2)\| \leq \frac{\|y_1 - y_2\|}{1 - \lambda}, \quad (17.191)$$

c'est-à-dire que

$$f^{-1}(B(y, r)) \subset B(f^{-1}(y), \frac{r}{1 - \lambda}), \quad (17.192)$$



ce qui signifie que  $f^{-1}$  est Lipschitz de constante souhaitée et donc continue.

Il reste à prouver que  $f(\mathcal{O})$  est ouvert. Pour cela nous prenons  $y_0 = f(x_0)$  dans  $f(\mathcal{O})$  est nous prouvons qu'il existe  $\epsilon$  tel que  $B(y_0, \epsilon)$  soit dans  $f(\mathcal{O})$ . Il faut donc que pour tout  $y \in B(y_0, \epsilon)$ , l'équation  $f(x) = y$  ait une solution. Nous considérons l'application

$$L_y: x \mapsto y - \varphi(x). \quad (17.193)$$

Ce que nous cherchons est un point fixe de  $L_y$  parce que si  $L_y(x) = x$  alors  $y = x + \varphi(x) = f(x)$ . Puisque

$$\|L_y(x) - L_y(x')\| = \|\varphi(x) - \varphi(x')\| \leq \lambda \|x - x'\|, \quad (17.194)$$

l'application  $L_y$  est une contraction de constante  $\lambda$ . Par ailleurs  $x_0$  est un point fixe de  $L_{y_0}$ , donc en vertu de la caractérisation (12.929) des fonctions Lipschitziennes,

$$L_{y_0}(\overline{B(x_0, \delta)}) \subset \overline{B(L_{y_0}(x_0), \lambda\delta)} = \overline{B(x_0, \lambda\delta)}. \quad (17.195)$$

Comme pour tout  $y$  et  $x$  nous avons  $L_y(x) = L_{y_0}(x) + y - y_0$ ,

$$L_y(\overline{B(x_0, \delta)}) = L_{y_0}(\overline{B(x_0, \delta)}) + (y - y_0) \subset \overline{B(x_0, \lambda\delta)} + (y - y_0) \subset \overline{B(x_0, \lambda\delta) + \|y - y_0\|}. \quad (17.196)$$

Si  $\epsilon < (1 - \lambda)\delta$  alors  $\lambda\delta + \|y - y_0\| < \delta$ . Un tel choix de  $\epsilon > 0$  est possible parce que  $\lambda < 1$ . Pour une telle valeur de  $\epsilon$  nous avons

$$L_y(\overline{B(x_0, \delta)}) \subset \overline{B(x_0, \delta)}. \quad (17.197)$$

Par conséquent  $L_y$  est une contraction sur l'espace métrique complet  $\overline{B(x_0, \delta)}$ , ce qui signifie que  $L_y$  y possède un point fixe par le théorème de Picard 17.34.  $\square$

Nous allons le démontrer dans le cas un peu plus général (mais pas plus cher<sup>26</sup>) des espaces de Banach en tant que conséquence du théorème de point fixe de Picard 17.34.

ThoXWpzqCn

**Théorème 17.47** (Inversion locale dans un espace de Banach[570, 404]).

Soit une fonction  $f \in C^p(E, F)$  avec  $p \geq 1$  entre deux espaces de Banach. Soit  $x_0 \in E$  tel que  $df_{x_0}$  soit une bijection bicontinue<sup>27</sup>. Alors il existe un voisinage ouvert  $V$  de  $x_0$  et  $W$  de  $f(x_0)$  tels que

- (1)  $f: V \rightarrow W$  soit une bijection,
- (2)  $f^{-1}: W \rightarrow V$  soit de classe  $C^p$ .

*Démonstration.* Nous commençons par simplifier un peu le problème. Pour cela, nous considérons la translation  $T: x \mapsto x + x_0$  et l'application linéaire

$$L: \mathbb{R}^n \rightarrow \mathbb{R}^n \quad (17.198)$$

$$x \mapsto (df_{x_0})^{-1}x$$

qui sont tous deux des difféomorphismes ( $L$  en est un par hypothèse d'inversibilité). Quitte à travailler avec la fonction  $k = L \circ f \circ T$ , nous pouvons supposer que  $x_0 = 0$  et que  $df_{x_0} = \mathbb{1}$ . Pour comprendre cela il faut utiliser deux fois la formule de différentielle de fonction composée de la proposition 12.792 :

$$dk_0(u) = dL_{(f \circ T)(0)}(df_{T(0)}dT_0(u)). \quad (17.199)$$

Puisque  $L$  est linéaire, sa différentielle est elle-même, c'est-à-dire  $dL_{(f \circ T)(0)} = (df_{x_0})^{-1}$ , et par ailleurs  $dT_0 = \mathbb{1}$ , donc

$$dk_0(u) = (df_{x_0})^{-1}(df_{x_0}(u)) = u, \quad (17.200)$$

ce qui signifie bien que  $dk_0 = \mathbb{1}$ . Pour tout cela nous avons utilisé en plein le fait que  $df_{x_0}$  était inversible.

26. Sauf la justification de la régularité de l'application  $A \mapsto A^{-1}$

27. En dimension finie, une application linéaire est toujours continue et d'inverse continu.

Nous posons  $g = f - \mathbb{1}$ , c'est-à-dire  $g(x) = f(x) - x$ , qui a la propriété  $dg_0 = 0$ . Étant donné que  $g$  est de classe  $C^1$ , l'application<sup>28</sup>

$$\begin{aligned} dg: E &\rightarrow \text{GL}(F) \\ x &\mapsto dg_x \end{aligned} \quad (17.201)$$

est continue. En conséquence, nous avons un voisinage  $U'$  de 0 pour lequel

$$\sup_{x \in U'} \|dg_x\| < \frac{1}{2}. \quad (17.202) \quad \text{EqSGTOfyx}$$

Maintenant le théorème des accroissements finis 10.324 (12.340 pour la dimension finie) nous indique que pour tout  $x, x' \in U'$  nous avons<sup>29</sup>

$$\|g(x') - g(x)\| \leq \sup_{a \in [x, x']} \|dg_a\| \cdot \|x - x'\| \leq \frac{1}{2} \|x - x'\|, \quad (17.203)$$

ce qui prouve que  $g$  est une contraction au moins sur l'ouvert  $U'$ . Nous allons aussi donner une idée de la façon dont  $f$  fonctionne : si  $x_1, x_2 \in U'$  alors

$$\|x_1 - x_2\| = \|g(x_1) - f(x_1) - g(x_2) + f(x_2)\| \quad (17.204a)$$

$$\leq \|g(x_1) - g(x_2)\| + \|f(x_1) - f(x_2)\| \quad (17.204b)$$

$$\leq \frac{1}{2} \|x_1 - x_2\| + \|f(x_1) - f(x_2)\|, \quad (17.204c)$$

ce qui montre que

$$\|x_1 - x_2\| \leq 2\|f(x_1) - f(x_2)\|. \quad (17.205)$$

Maintenant que nous savons que  $g$  est contractante de constante  $\frac{1}{2}$  et que  $f = g + \mathbb{1}$  nous pouvons utiliser la proposition 17.46 pour conclure que  $f$  est un homéomorphisme sur un ouvert  $U$  (partie de  $U'$ ) de  $E$  et  $f^{-1}$  a une constante de Lipschitz plus petite ou égale à  $(1 - \frac{1}{2})^{-1} = 2$ .

Nous allons maintenant prouver que  $f^{-1}$  est différentiable et que sa différentielle est donnée par  $(df^{-1})_{f(x)} = (df_x)^{-1}$ .

Soient  $a, b \in U$  et  $u = b - a$ . Étant donné que  $f$  est différentiable en  $a$ , il existe une fonction  $\alpha \in o(\|u\|)$  telle que

$$f(b) - f(a) - df_a(u) = \alpha(u). \quad (17.206)$$

En notant  $y_a = f(a)$  et  $y_b = f(b)$  et en appliquant  $(df_a)^{-1}$  à cette dernière équation,

$$(df_a)^{-1}(y_b - y_a) - u = (df_a)^{-1}(\alpha(u)). \quad (17.207)$$

Puisque  $df_a$  est bornée (et son inverse aussi), le membre de droite est encore une fonction  $\beta$  ayant la propriété  $\lim_{u \rightarrow 0} \beta(u)/\|u\| = 0$ ; en réordonnant les termes,

$$b - a = (df_a)^{-1}(y_b - y_a) + \beta(u) \quad (17.208)$$

et donc

$$f^{-1}(y_b) - f^{-1}(y_a) - (df_a)^{-1}(y_b - y_a) = \beta(u), \quad (17.209)$$

ce qui prouve que  $f^{-1}$  est différentiable et que  $(df^{-1})_{y_a} = (df_a)^{-1}$ .

La différentielle  $df^{-1}$  est donc obtenue par la chaîne

$$df^{-1}: f(U) \xrightarrow{f^{-1}} U' \xrightarrow{df} \text{GL}(F) \xrightarrow{\text{Inv}} \text{GL}(F) \quad (17.210)$$

28. Ici  $\text{GL}(F)$  est l'ensemble des applications linéaires, inversibles et continues de  $F$  dans lui-même. Ce ne sont pas spécialement des matrices parce que nous n'avons pas d'hypothèses sur la dimension de  $F$ , finie ou non.

29. Ici nous supposons avoir choisi  $U'$  convexe afin que tous les  $a \in [x, x']$  soient bien dans  $U'$  et donc soumis à l'inéquation (17.202), ce qui est toujours possible, il suffit de prendre une boule.

où l'application  $\text{Inv}: \text{GL}(F) \rightarrow \text{GL}(F)$  est l'application  $X \mapsto X^{-1}$  qui est de classe  $C^\infty$  par le théorème 10.332. D'autre part, par hypothèse  $df$  est une application de classe  $C^{k-1}$  et donc au minimum  $C^0$  parce que  $k \geq 1$ . Enfin, l'application  $f^{-1}: f(U) \rightarrow U$  est continue (parce que la proposition 17.46 précise que  $f$  est un homéomorphisme). Donc toute la chaîne est continue et  $df^{-1}$  est continue. Cela entraîne immédiatement que  $f^{-1}$  est  $C^1$  et donc que toute la chaîne est  $C^1$ .

Par récurrence nous obtenons la chaîne

$$df^{-1}: f(U) \xrightarrow[C^{k-1}]{f^{-1}} U' \xrightarrow[C^{k-1}]{df} \text{GL}(F) \xrightarrow[C^\infty]{\text{Inv}} \text{GL}(F) \quad (17.211)$$

qui prouve que  $df^{-1}$  est  $C^{k-1}$  et donc que  $f^{-1}$  est  $C^k$ . La récurrence s'arrête ici parce que  $df$  n'est pas mieux que  $C^{k-1}$ .  $\square$

### 17.8.3 Théorème de la fonction implicite

Nous énonçons et démontrons le théorème de la fonction implicite dans le cas d'espaces de Banach.

ThoAcaWho

**Théorème 17.48** (Théorème de la fonction implicite dans Banach[406]).

Soient  $E$ ,  $F$  et  $G$  des espaces de Banach et des ouverts  $U \subset E$ ,  $V \subset F$ . Nous considérons une fonction  $f: U \times V \rightarrow G$  de classe  $C^r$  telle que<sup>30</sup>

$$d_y f_{(x_0, y_0)}: F \rightarrow G \quad (17.212)$$

soit un isomorphisme pour un certain  $(x_0, y_0) \in U \times V$ .

Alors nous avons des voisinages  $U_0$  de  $x_0$  dans  $E$  et  $W_0$  de  $f(x_0, y_0)$  dans  $G$  et une fonction de classe  $C^r$

$$g: U_0 \times W_0 \rightarrow V \quad (17.213)$$

telle que

$$f(x, g(x, w)) = w \quad (17.214)$$

pour tout  $(x, w) \in U_0 \times W_0$ .

Cette fonction  $g$  est unique au sens suivant : il existe un voisinage  $V_0$  de  $y_0$  tel que si  $(x, y) \in U_0 \times V_0$  et  $w \in W_0$  satisfont à  $f(x, y) = w$  alors  $y = g(x, w)$ . Autrement dit, la fonction  $g: U_0 \times W_0 \rightarrow V_0$  est unique.

*Démonstration.* Nous commençons par considérer la fonction

$$\begin{aligned} \Phi: U \times V &\rightarrow E \times G \\ (x, y) &\mapsto (x, f(x, y)) \end{aligned} \quad (17.215)$$

et sa différentielle

$$d\Phi_{(x_0, y_0)}(u, v) = \frac{d}{dt} \left[ (x_0 + tu, f(x_0 + tu, y_0 + tv)) \right]_{t=0} \quad (17.216a)$$

$$= \left( \frac{d}{dt} [x_0 + tu]_{t=0}, \frac{d}{dt} [f(x_0 + tu, y_0 + tv)]_{t=0} \right) \quad (17.216b)$$

$$= (u, df_{(x_0, y_0)}(u, v)). \quad (17.216c)$$

Nous utilisons alors la proposition 10.330 pour conclure que

$$d\Phi_{(x_0, y_0)}(u, v) = (u, (d_1 f)_{(x_0, y_0)}(u) + (d_2 f)_{(x_0, y_0)}(v)), \quad (17.217)$$

mais comme par hypothèse  $(d_2 f)_{(x_0, y_0)}: F \rightarrow G$  est un isomorphisme, l'application  $d\Phi_{(x_0, y_0)}: E \times F \rightarrow E \times G$  est également un isomorphisme. Par conséquent le théorème d'inversion locale 17.47

30. La notation  $d_y$  est la différentielle partielle de la définition 10.329.

nous indique qu'il existe un voisinage  $\mathcal{O}$  de  $(x_0, y_0)$  et  $\mathcal{P}$  de  $\Phi(x_0, y_0)$  tels que  $\Phi: \mathcal{O} \rightarrow \mathcal{P}$  soit une bijection et  $\Phi^{-1}: \mathcal{P} \rightarrow \mathcal{O}$  soit de classe  $C^r$ . Comme  $\mathcal{P}$  est un voisinage de

$$\Phi(x_0, y_0) = (x_0, f(x_0, y_0)), \quad (17.218)$$

nous pouvons par 7.285 le choisir un peu plus petit de telle sorte à avoir  $\mathcal{P} = U_0 \times W_0$  où  $U_0$  est un voisinage de  $x_0$  et  $W_0$  un voisinage de  $f(x_0, y_0)$ . Dans ce cas nous devons obligatoirement aussi restreindre  $\mathcal{O}$  à  $U_0 \times V_0$  pour un certain voisinage  $V_0$  de  $y_0$ . L'application  $\Phi^{-1}$  a obligatoirement la forme

$$\begin{aligned} \Phi^{-1}: U_0 \times W_0 &\rightarrow U_0 \times V_0 \\ (x, w) &\mapsto (x, g(x, w)) \end{aligned} \quad \text{EqMHT_QzHRn} \quad (17.219)$$

pour une certaine fonction  $g: U_0 \times W_0 \rightarrow V$ . Cette fonction  $g$  est la fonction cherchée parce qu'en appliquant  $\Phi$  à (17.219),

$$(x, w) = \Phi(x, g(x, w)) = (x, f(x, g(x, w))), \quad (17.220)$$

qui nous dit que pour tout  $x \in U_0$  et tout  $w \in W_0$  nous avons

$$f(x, g(x, w)) = w. \quad (17.221)$$

Si vous avez bien suivi le sens de l'équation (17.219) alors vous avez compris l'unicité. Sinon, considérez  $(x, y) \in U_0 \times V_0$  et  $w \in W_0$  tels que  $f(x, y) = w$ . Alors  $(x, f(x, y)) = (x, w)$  et

$$\Phi(x, y) = (x, w). \quad (17.222)$$

Mais vu que  $\Phi: U_0 \times V_0 \rightarrow U_0 \times W_0$  est une bijection, cette relation définit de façon univoque l'élément  $(x, y)$  de  $U_0 \times V_0$ , qui ne sera autre que  $g(x, w)$ .  $\square$

Le théorème de la fonction implicite s'énonce de la façon suivante pour des espaces de dimension finie.

ThoRYN\_jvZrZ

**Théorème 17.49** (Théorème de la fonction implicite en dimension finie).

Soit une fonction  $F: \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$  de classe  $C^k$  et  $(\alpha, \beta) \in \mathbb{R}^n \times \mathbb{R}^m$  tels que

$$(1) \quad F(\alpha, \beta) = 0,$$

$$(2) \quad \frac{\partial(F_1, \dots, F_m)}{\partial(y_1, \dots, y_m)} \neq 0, \text{ c'est-à-dire que } (d_y F)_{(\alpha, \beta)} \text{ est inversible.}$$

Alors il existe un voisinage ouvert  $V$  de  $\alpha$  dans  $\mathbb{R}^n$ , un voisinage ouvert  $W$  de  $\beta$  dans  $\mathbb{R}^m$  et une application  $\varphi: V \rightarrow W$  de classe  $C^k$  telle que pour tout  $x \in V$  on ait

$$F(x, \varphi(x)) = 0. \quad (17.223)$$

De plus si  $(x, y) \in V \times W$  satisfait à  $F(x, y) = 0$ , alors  $y = \varphi(x)$ .

RemPYA\_pkTEx

**Remarque 17.50.**

Notons que cet énoncé est tourné un peu différemment en ce qui concerne le nombre de variables dont dépend la fonction implicite : comparez

$$f(x, g(x, w)) = w \quad (17.224a)$$

$$F(x, \varphi(x)) = 0. \quad (17.224b)$$

Le deuxième est un cas particulier du premier en posant

$$F(x, y) = f(x, y) - f(x_0, y_0) \quad (17.225)$$

et donc en considérant  $w$  comme valant la constante  $f(x_0, y_0)$  ; dans ce cas la fonction  $g$  ne dépend plus que de la variable  $x$ .

**Exemple 17.51.**

La remarque 17.50 signifie entre autres que le théorème 17.48 est plus fort que 17.49 parce que le premier permet de choisir la valeur d'arrivée. Parlons de l'exemple classique du cercle et de la fonction  $f(x, y) = x^2 + y^2$ . Nous savons que

$$f(\alpha, \beta) = 1. \quad (17.226)$$

Alors le théorème 17.48 nous donne une fonction  $g$  telle que

$$f(x, g(x, r)) = r \quad (17.227)$$

tant que  $x$  est proche de  $\alpha$ , que  $r$  est proche de 1 et que  $g$  donne des valeurs proches de  $\beta$ .

L'énoncé 17.49 nous oblige à travailler avec la fonction  $F(x, y) = x^2 + y^2 - 1$ , de telle sorte que

$$F(\alpha, \beta) = 0, \quad (17.228)$$

et que nous ayons une fonction  $\varphi$  telle que

$$F(x, \varphi(x)) = 0. \quad (17.229)$$

La fonction  $\varphi$  ne permet donc que de trouver des points sur le cercle de rayon 1.  $\triangle$

**17.8.4 Exemple**

Le théorème de la fonction implicite a pour objet de donner l'existence de la fonction  $\varphi$ . Maintenant nous pouvons dire beaucoup de choses sur les dérivées de  $\varphi$  en considérant la fonction

$$x \mapsto F(x, \varphi(x)). \quad (17.230)$$

Par définition de  $\varphi$ , cette fonction est toujours nulle. En particulier, nous pouvons dériver l'équation

$$F(x, \varphi(x)) = 0, \quad (17.231)$$

et nous trouvons plein de choses.

EXooTLNAooCJHPnq

**Exemple 17.52.**

Prenons par exemple la fonction <sup>31</sup>

$$F((x, y), z) = ze^z - x - y, \quad (17.232)$$

Le théorème de la fonction implicite 17.48 nous permet de considérer la fonction  $z(x, y)$  vérifiant

$$F(x, y, z(x, y)) = 0, \quad (17.233)$$

c'est-à-dire telle que

$$z(x, y)e^{z(x, y)} - x - y = 0. \quad (17.234)$$

EqDefZImplExemple

pour tout  $x$  et  $y \in \mathbb{R}$ . Que pouvons dire de la fonction  $z$  ?

Nous pouvons facilement trouver  $z(0, 0)$  parce que

$$z(0, 0)e^{z(0, 0)} = 0, \quad (17.235)$$

donc  $z(0, 0) = 0$ .

Nous pouvons dire des choses sur les dérivées de  $z(x, y)$ . Voyons par exemple  $(\partial_x z)(x, y)$ . Pour trouver cette dérivée, nous dérivons la relation (17.234) par rapport à  $x$ . Ce que nous trouvons est

$$(\partial_x z)e^z + ze^z(\partial_x z) - 1 = 0. \quad (17.236)$$

---

31. Définition de l'exponentielle : 15.66.

Cette équation peut être résolue par rapport à  $\partial_x z$  :

$$\frac{\partial z}{\partial x}(x, y) = \frac{1}{e^z(1+z)}. \quad (17.237)$$

Remarquez que cette équation ne donne pas tout à fait la dérivée de  $z$  en fonction de  $x$  et  $y$ , parce que  $z$  apparaît dans l'expression, alors que  $z$  est justement la fonction inconnue. En général, c'est la vie, nous ne pouvons pas faire mieux.

Dans certains cas, on peut aller plus loin. Par exemple, nous pouvons calculer cette dérivée au point  $(x, y) = (0, 0)$  parce que  $z(0, 0)$  est connu :

$$\frac{\partial z}{\partial x}(0, 0) = 1. \quad (17.238)$$

Ceci est pratique pour calculer, par exemple, le développement en Taylor de  $z$  autour de  $(0, 0)$ .  $\triangle$

### Exemple 17.53.

Est-ce que l'équation  $e^y + xy = 0$  définit au moins localement une fonction  $y(x)$  ? Nous considérons la fonction

$$f(x, y) = \begin{pmatrix} x \\ e^y + xy \end{pmatrix} \quad (17.239)$$

La différentielle de cette application est

$$df_{(0,0)}(u) = \frac{d}{dt} [f(tu_1, tu_2)]_{t=0} = \frac{d}{dt} \begin{pmatrix} tu_1 \\ e^{tu_2} + t^2 u_1 u_2 \end{pmatrix}_{t=0} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \quad (17.240)$$

L'application  $f$  définit donc un difféomorphisme local autour des points  $(x_0, y_0)$  et  $f(x_0, y_0)$ . Soit  $(u, 0)$  un point dans le voisinage de  $f(x_0, y_0)$ . Alors il existe un unique  $(x, y)$  tel que

$$f(x, y) = \begin{pmatrix} x \\ e^y + xy \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix}. \quad (17.241)$$

Nous avons automatiquement  $x = u$  et  $e^y + xy = 0$ . Notons toutefois que pour que ce procédé donne effectivement une fonction implicite  $y(x)$  nous devons avoir des points de la forme  $(u, 0)$  dans le voisinage de  $f(x_0, y_0)$ .  $\triangle$

## 17.9 Décomposition polaire (régularité)

### 17.54.

Nous allons montrer que l'application

$$\begin{aligned} f: S^{++}(n, \mathbb{R}) &\rightarrow S^{++}(n, \mathbb{R}) \\ A &\mapsto \sqrt{A} \end{aligned} \quad (17.242)$$

est un difféomorphisme.

Cependant  $S^{++}(n, \mathbb{R})$  n'est pas un ouvert de  $\mathbb{M}(n, \mathbb{R})$  et nous ne savons pas ce qu'est la différentielle d'une application non définie sur un ouvert. Nous allons donc en réalité montrer que l'application racine carrée existe sur un voisinage de chacun des points de  $S^{++}(n, \mathbb{R})$ . Et comme une union quelconque d'ouverts est un ouvert, la fonction  $f$  sera bien définie sur un ouvert de  $\mathbb{M}(n, \mathbb{R})$ .

### Lemme 17.55.

L'application

$$\begin{aligned} f: S^{++}(n, \mathbb{R}) &\rightarrow S^{++}(n, \mathbb{R}) \\ A &\mapsto A^2 \end{aligned} \quad (17.243)$$

est un  $C^\infty$ -difféomorphisme.

NomDJMUooTRUVkS

LemLBFOooDdNcgy

*Démonstration.* Prouvons d'abord que  $f$  prend ses valeurs dans  $S^{++}(n, \mathbb{R})$ . Si  $A \in S^{++}(n, \mathbb{R})$  alors par la diagonalisation 8.226 elle s'écrit  $A = QDQ^{-1}$  où  $D$  est diagonale avec des nombres strictement positifs sur la diagonale. Avec cela,  $A^2 = QD^2Q^{-1}$  où  $D^2$  contient encore des nombres strictement positifs sur la diagonale.

L'application  $f$  étant essentiellement des polynômes en les entrées de  $A$ , elle est de classe  $C^\infty$ .

Passons à l'étude de la différentielle. Comme mentionné en 17.54 nous allons en réalité voir  $f$  sur un ouvert de  $\mathbb{M}(n, \mathbb{R})$  autour de  $A \in S^{++}(n, \mathbb{R})$ . Par conséquent si  $A \in S^{++}(n, \mathbb{R})$ ,

$$df: S^{++}(n, \mathbb{R}) \rightarrow \mathcal{L}(\mathbb{M}(n, \mathbb{R}), \mathbb{M}(n, \mathbb{R})) \quad (17.244a)$$

$$df_A: \mathbb{M}(n, \mathbb{R}) \rightarrow \mathbb{M}(n, \mathbb{R}). \quad (17.244b)$$

Le calcul de  $df_A$  est facile. Soit  $u \in \mathbb{M}(n, \mathbb{R})$  et faisons le calcul en utilisant la formule du lemme (12.280) :

$$df_A(u) = \frac{d}{dt} \left[ f(A + tu) \right]_{t=0} \quad (17.245a)$$

$$= \frac{d}{dt} \left[ A^2 + tAu + tuA + t^2u^2 \right]_{t=0} \quad (17.245b)$$

$$= Au + uA. \quad (17.245c)$$

Nous allons utiliser le théorème d'inversion locale 17.47 à la fonction  $f$ . Dans la suite,  $A$  est une matrice de  $S^{++}(n, \mathbb{R})$ .

- (1)  **$df_A$  est injective** Soit  $M \in \mathbb{M}(n, \mathbb{R})$  dans le noyau de  $df_A$ . En posant  $M' = Q^{-1}MQ$  nous avons  $M = QM'Q^{-1}$  et on applique  $df_A$  à  $QM'Q^{-1}$  :

$$df_A(QM'Q^{-1}) = Q(DM + MD)Q^{-1}. \quad (17.246)$$

où  $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$  avec  $\lambda_i > 0$ . La matrice  $D$  est inversible. Nous avons  $M' = -DM'D^{-1}$ , et en coordonnées,

$$M'_{ij} = - \sum_{kl} D_{ik} M'_{kl} D_{lj}^{-1} \quad (17.247a)$$

$$= - \sum_{kl} \lambda_i \delta_{ik} M'_{kl} \frac{1}{\lambda_j} \delta_{lj} \quad (17.247b)$$

$$= - \frac{\lambda_i}{\lambda_j} M'_{ij}. \quad (17.247c)$$

C'est-à-dire que  $M'_{ij} = -\frac{\lambda_i}{\lambda_j} M'_{ij}$  avec  $-\frac{\lambda_i}{\lambda_j} < 0$ . Cela implique  $M' = 0$  et par conséquent  $M = 0$ .

- (2)  **$df_A$  est surjective** Soit  $N \in \mathbb{M}(n, \mathbb{R})$ ; nous cherchons  $M \in \mathbb{M}(n, \mathbb{R})$  tel que  $df_A(M) = N$ . Nous posons  $N' = Q^{-1}NQ$  et  $M = QM'Q^{-1}$ , ce qui nous donne à résoudre  $df_D(M') = N'$ . Passons en coordonnées :

$$(DM' + M'D)_{ij} = \sum_k (\delta_{ik} \lambda_i M'_{kj} + M'_{ik} \delta_{kj} \lambda_j) = M'_{ij} (\lambda_i + \lambda_j) \quad (17.248)$$

où  $\lambda_i + \lambda_j \neq 0$ . Il suffit donc de prendre la matrice  $M'$  donnée par

$$M'_{ij} = \frac{1}{\lambda_i + \lambda_j} N'_{ij} \quad (17.249)$$

pour que  $df_A(M') = N'$ .

Le théorème d'inversion locale donne un voisinage  $V$  de  $A$  dans  $\mathbb{M}(n, \mathbb{R})$  et un voisinage  $W$  de  $A^2$  dans  $\mathbb{M}(n, \mathbb{R})$  tels que  $f: V \rightarrow W$  soit une bijection et  $f^{-1}: W \rightarrow V$  soit de même régularité, en l'occurrence  $C^\infty$ .  $\square$

**Remarque 17.56.**

Oui, il y a des matrices non symétriques qui ont une unique racine carrée.

La proposition suivante, qui dépend du théorème d'inversion locale par le lemme 17.55, donne plus de régularité à la décomposition polaire donnée dans le théorème 13.38.

PropWCXAooDuFMjn

**Proposition 17.57** (Décomposition polaire : cas réel (suite)).

*L'application*

$$\begin{aligned} f: \mathcal{O}(n, \mathbb{R}) \times S^{++}(n, \mathbb{R}) &\rightarrow \mathrm{GL}(n, \mathbb{R}) \\ (Q, S) &\mapsto SQ \end{aligned} \quad (17.250)$$

est un difféomorphisme de classe  $C^\infty$ .

*Démonstration.* Si  $M$  est donnée dans  $\mathrm{GL}(n, \mathbb{R})$  alors la décomposition polaire<sup>32</sup>  $M = QS$  est donnée par  $S = \sqrt{MM^t}$  et  $Q = MS^{-1}$ . Autrement dit, si nous considérons la fonction de décomposition polaire

$$f: \mathcal{O}(n, \mathbb{R}) \times S^{++}(n, \mathbb{R}) \rightarrow \mathrm{GL}(n, \mathbb{R}) \quad (17.251)$$

alors

$$f^{-1}(M) = (M(\sqrt{MM^t})^{-1}, \sqrt{MM^t}). \quad (17.252)$$

Nous avons vu dans le lemme 17.55 que la racine carrée était un  $C^\infty$ -difféomorphisme. Le reste n'étant que des produits de matrices, la régularité est de mise.  $\square$

## 17.10 Théorème de Von Neumann

**Lemme 17.58** ([126]).

Soit  $G$ , un sous-groupe fermé de  $\mathrm{GL}(n, \mathbb{R})$  et

$$\mathcal{L}_G = \{m \in \mathbb{M}(n, \mathbb{R}) \text{ tel que } e^{tm} \in G \forall t \in \mathbb{R}\}. \quad (17.253)$$

Alors  $\mathcal{L}_G$  est un sous-espace vectoriel de  $\mathbb{M}(n, \mathbb{R})$ .

*Démonstration.* Si  $m \in \mathcal{L}_G$ , alors  $\lambda m \in \mathcal{L}_G$  par construction. Le point délicat à prouver est le fait que si  $a, b \in \mathcal{L}_G$ , alors  $a + b \in \mathcal{L}_G$ . Soit  $a \in \mathbb{M}(n, \mathbb{R})$ ; nous savons qu'il existe une fonction  $\alpha_a: \mathbb{R} \rightarrow \mathbb{M}$  telle que

$$e^{ta} = \mathbb{1} + ta + \alpha_a(t) \quad (17.254)$$

et

$$\lim_{t \rightarrow 0} \frac{\alpha(t)}{t} = 0. \quad (17.255)$$

Si  $a$  et  $b$  sont dans  $\mathcal{L}_G$ , alors  $e^{ta}e^{tb} \in G$ , mais il n'est pas vrai en général que cela soit égal à  $e^{t(a+b)}$ . Pour tout  $k \in \mathbb{N}$  nous avons

$$e^{a/k}e^{b/k} = \left(\mathbb{1} + \frac{a}{k} + \alpha_a\left(\frac{1}{k}\right)\right) \left(\mathbb{1} + \frac{b}{k} + \alpha_b\left(\frac{1}{k}\right)\right) = \mathbb{1} + \frac{a+b}{k} + \beta\left(\frac{1}{k}\right) \quad (17.256)$$

où  $\beta: \mathbb{R} \rightarrow \mathbb{M}$  est encore une fonction vérifiant  $\beta(t)/t \rightarrow 0$ . Si  $k$  est assez grand, nous avons

$$\left\| \frac{a+b}{k} + \beta\left(\frac{1}{k}\right) \right\| < 1, \quad (17.257)$$

---

32. Proposition 13.38.



et nous pouvons profiter du lemme 15.166 pour écrire alors

$$\left(e^{a/k} e^{b/k}\right)^k = e^{k \ln \left(1 + \frac{a+b}{k} + \beta\left(\frac{1}{k}\right)\right)}. \quad (17.258)$$

Ce qui se trouve dans l'exponentielle est

$$k \left[ \frac{a+b}{k} + \alpha\left(\frac{1}{k}\right) + \sigma \left( \frac{a+b}{k} + \alpha\left(\frac{1}{k}\right) \right) \right]. \quad (17.259)$$

Les diverses propriétés vues montrent que le tout tend vers  $a+b$  lorsque  $k \rightarrow \infty$ . Par conséquent

$$\lim_{k \rightarrow \infty} \left(e^{a/k} e^{b/k}\right)^k = e^{a+b}. \quad (17.260)$$

Ce que nous avons prouvé est que pour tout  $t$ ,  $e^{t(a+b)}$  est une limite d'éléments dans  $G$  et est donc dans  $G$  parce que ce dernier est fermé.  $\square$

Comme  $\mathcal{L}_G$  est un sous-espace vectoriel de  $\mathbb{M}(n, \mathbb{R})$ , nous pouvons considérer un supplémentaire  $M$ .

LemHOsbREC

### Lemme 17.59.

Il n'existe pas de suite  $(m_k)$  dans  $M \setminus \{0\}$  convergeant vers zéro et telle que  $e^{m_k} \in G$  pour tout  $k$ .

*Démonstration.* Supposons que nous ayons  $m_k \rightarrow 0$  dans  $M \setminus \{0\}$  avec  $e^{m_k} \in G$ . Nous considérons les éléments  $\epsilon_k = \frac{m_k}{\|m_k\|}$  qui sont sur la sphère unité de  $\mathbb{GL}(n, \mathbb{R})$ . Quitte à prendre une sous-suite, nous pouvons supposer que cette suite converge, et puisque  $M$  est fermé, ce sera vers  $\epsilon \in M$  avec  $\|\epsilon\| = 1$ . Pour tout  $t \in \mathbb{R}$  nous avons

$$e^{t\epsilon} = \lim_{k \rightarrow \infty} e^{t\epsilon_k}. \quad (17.261)$$

En vertu de la décomposition d'un réel en partie entière et décimale, pour tout  $k$  nous avons  $\lambda_k \in \mathbb{Z}$  et  $|\mu_k| \leq \frac{1}{2}$  tel que  $t/\|m_k\| = \lambda_k + \mu_k$ . Avec ça,

$$e^{t\epsilon} = \lim_{k \rightarrow \infty} \exp\left(\frac{t}{m_k} m_k\right) = \lim_{k \rightarrow \infty} e^{\lambda_k m_k} e^{\mu_k m_k}. \quad (17.262)$$

Pour tout  $k$  nous avons  $e^{\lambda_k m_k} \in G$ . De plus  $|\mu_k|$  étant borné et  $m_k$  tendant vers zéro nous avons  $e^{\mu_k m_k} \rightarrow 1$ . Au final

$$e^{t\epsilon} = \lim_{k \rightarrow \infty} e^{t\epsilon_k} \in G \quad (17.263)$$

Cela signifie que  $\epsilon \in \mathcal{L}_G$ , ce qui est impossible parce que nous avons déjà dit que  $\epsilon \in M \setminus \{0\}$ .  $\square$

LemGGTtxdF

### Lemme 17.60.

*L'application*

$$\begin{aligned} f: \mathcal{L}_G \times M &\rightarrow \mathbb{GL}(n, \mathbb{R}) \\ l, m &\mapsto e^l e^m \end{aligned} \quad (17.264)$$

est un difféomorphisme local entre un voisinage de  $(0, 0)$  dans  $\mathbb{M}(n, \mathbb{R})$  et un voisinage de  $1$  dans  $\exp(\mathbb{M}(n, \mathbb{R}))$ .

Notons que nous ne disons rien de  $e^{\mathbb{M}(n, \mathbb{R})}$ . Nous n'allons pas nous embarquer à discuter si ce serait tout  $\mathbb{GL}(n, \mathbb{R})$  <sup>33</sup> ou bien si ça contiendrait ne fut-ce que  $G$ .

*Démonstration.* Le fait que  $f$  prenne ses valeurs dans  $\mathbb{GL}(n, \mathbb{R})$  est simplement dû au fait que les exponentielles sont toujours inversibles. Nous considérons ensuite la différentielle : si  $u \in \mathcal{L}_G$  et  $v \in M$  nous avons

$$df_{(0,0)}(u, v) = \frac{d}{dt} \left[ f(t(u, v)) \right]_{t=0} = \frac{d}{dt} \left[ e^{tu} e^{tv} \right]_{t=0} = u + v. \quad (17.265)$$

33. Vu les dimensions y'a tout de même peu de chance.

L'application  $df_0$  est donc une bijection entre  $\mathcal{L}_G \times M$  et  $\mathbb{M}(n, \mathbb{R})$ . Le théorème d'inversion locale 17.47 nous assure alors que  $f$  est une bijection entre un voisinage de  $(0, 0)$  dans  $\mathcal{L}_G \times M$  et son image. Mais comme  $df_0$  est une bijection avec  $\mathbb{M}(n, \mathbb{R})$ , l'image en question contient un ouvert autour de  $\mathbb{1}$  dans  $\exp(\mathbb{M}(n, \mathbb{R}))$ .  $\square$

ThoOBriEoe

**Théorème 17.61** (Von Neumann[126, 571, 572]).

*Tout sous-groupe fermé de  $\mathrm{GL}(n, \mathbb{R})$  est une sous-variété de  $\mathrm{GL}(n, \mathbb{R})$ .*

*Démonstration.* Soit  $G$  un tel groupe ; nous devons prouver que c'est localement difféomorphe à un ouvert de  $\mathbb{R}^n$ . Et si on est pervers, on ne va pas faire localement difféomorphe à un ouvert de  $\mathbb{R}^n$ , mais à un ouvert d'un espace vectoriel de dimension finie. Nous allons être pervers.

Étant donné que pour tout  $g \in G$ , l'application

$$\begin{aligned} L_g: G &\rightarrow G \\ h &\mapsto gh \end{aligned} \tag{17.266}$$

est de classe  $C^\infty$  et d'inverse  $C^\infty$ , il suffit de prouver le résultat pour un voisinage de  $\mathbb{1}$ .

Supposons d'abord que  $\mathcal{L}_G = \{0\}$ . Alors 0 est un point isolé de  $\ln(G)$  ; en effet si ce n'était pas le cas nous aurions un élément  $m_k$  de  $\ln(G)$  dans chaque boule  $B(0, r_k)$ . Nous aurions alors  $m_k = \ln(a_k)$  avec  $a_k \in G$  et donc

$$e^{m_k} = a_k \in G. \tag{17.267}$$

De plus  $m_k$  appartient forcément à  $M$  parce que  $\mathcal{L}_G$  est réduit à zéro. Cela nous donnerait une suite  $m_k \rightarrow 0$  dans  $M$  dont l'exponentielle reste dans  $G$ . Or cela est interdit par le lemme 17.59. Donc 0 est un point isolé de  $\ln(G)$ . L'application  $\ln$  étant continue<sup>34</sup>, nous en déduisons que  $\mathbb{1}$  est isolé dans  $G$ . Par le difféomorphisme  $L_g$ , tous les points de  $G$  sont isolés ; ce groupe est donc discret et par voie de conséquence, une variété.

Nous supposons maintenant que  $\mathcal{L}_G \neq \{0\}$ . Nous savons par la proposition 15.165 que

$$\exp: \mathbb{M}(n, \mathbb{R}) \rightarrow \mathbb{M}(n, \mathbb{R}) \tag{17.268}$$

est une application  $C^\infty$  vérifiant  $d\exp_0 = \mathrm{Id}$ . Nous pouvons donc utiliser le théorème d'inversion locale 17.47 qui nous offre donc l'existence d'un voisinage  $U$  de 0 dans  $\mathbb{M}(n, \mathbb{R})$  tel que  $W = \exp(U)$  soit un ouvert de  $\mathrm{GL}(n, \mathbb{R})$  et que  $\exp: U \rightarrow W$  soit un difféomorphisme de classe  $C^\infty$ .

Montrons que quitte à restreindre  $U$  (et donc  $W$  qui reste par définition l'image de  $U$  par  $\exp$ ), nous pouvons avoir  $\exp(U \cap \mathcal{L}_G) = W \cap G$ . D'abord  $\exp(\mathcal{L}_G) \subset G$  par construction. Nous avons donc  $\exp(U \cap \mathcal{L}_G) \subset W \cap G$ . Pour trouver une restriction de  $U$  pour laquelle nous avons l'égalité, nous supposons que pour tout ouvert  $\mathcal{O}$  dans  $U$ ,

$$\exp: \mathcal{O} \cap \mathcal{L}_G \rightarrow \exp(\mathcal{O}) \cap G \tag{17.269}$$

ne soit pas surjective. Cela donnerait un élément de  $\mathcal{O} \cap \mathcal{L}_G$  dont l'image par  $\exp$  n'est pas dans  $G$ . Nous construisons ainsi une suite en considérant une boule  $B(0, \frac{1}{k})$  incluse à  $U$  et  $x_k \in B(0, \frac{1}{k}) \cap \mathcal{L}_G$  vérifiant  $e^{x_k} \in G$ . D'après le choix des boules, nous avons évidemment  $x_k \rightarrow 0$ .

L'élément  $e^{x_k}$  est dans  $e^{\mathbb{M}(n, \mathbb{R})}$  et le difféomorphisme du lemme 17.60<sup>35</sup> nous donne  $(l_k, m_k) \in \mathcal{L}_G \times M$  tel que  $e^{l_k} e^{m_k} = e^{x_k}$ . À ce point nous considérons  $k$  suffisamment grand pour que  $e^{x_k}$  soit dans la partie de l'image de  $f$  sur lequel nous avons le difféomorphisme. Plus prosaïquement, nous posons

$$(l_k, m_k) = f^{-1}(e^{x_k}) \tag{17.270}$$

et nous profitons de la continuité pour permuter la limite avec  $f^{-1}$  :

$$\lim_{k \rightarrow \infty} (l_k, m_k) = f^{-1}\left(\lim_{k \rightarrow \infty} e^{x_k}\right) = f^{-1}(\mathbb{1}) = (0, 0). \tag{17.271}$$

34. Par le lemme 15.166.

35. Il me semble que l'utilisation de ce lemme manque à l'avant-dernière ligne de la preuve chez [126].

En particulier  $m_k \rightarrow 0$  alors que  $e^{m_k} = e^{x_k} e^{-l_k} \in G$ . La suite  $m_k$  viole le lemme 17.59. Nous pouvons donc restreindre  $U$  de telle façon à avoir

$$\exp(U \cap \mathcal{L}_G) = W \cap G. \quad (17.272)$$

Nous avons donc un ouvert de  $\mathcal{L}_G$  (l'ouvert  $U \cap \mathcal{L}_G$ ) qui est difféomorphe avec l'ouvert  $W \cap G$  de  $G$ . Donc  $G$  est une variété et accepte  $\mathcal{L}_G$  comme carte locale.  $\square$

### Remarque 17.62.

En termes savants, nous avons surtout montré que si  $G$  est un groupe de Lie d'algèbre de Lie  $\mathfrak{g}$ , alors l'exponentielle donne un difféomorphisme local entre  $\mathfrak{g}$  et  $G$ .

## 17.11 Recherche d'extrémums

DEFooJMMLOoYMZehc

### Définition 17.63 ([573]).

Si  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  est une fonction, le point  $a \in \mathbb{R}^n$  est un **maximum local** de  $f$  si il existe  $\delta > 0$  tel que pour tout  $x \in B(a, \delta)$ ,  $f(x) \leq f(a)$ .

Je vous laisse deviner la définition d'un **minimum local**.

Un **extrémum local** est un point qui est soit un minimum soit un maximum local.

#### 17.11.1 Extrema à une variable

### Définition 17.64.

Soit  $f: A \subset \mathbb{R} \rightarrow \mathbb{R}$  et  $a \in A$ . Le point  $a$  est un **maximum local** de  $f$  si il existe un voisinage  $\mathcal{U}$  de  $a$  tel que  $f(a) \geq f(x)$  pour tout  $x \in \mathcal{U} \cap A$ . Le point  $a$  est un **maximum global** si  $f(a) \geq g(x)$  pour tout  $x \in A$ .

La proposition basique à utiliser lors de la recherche d'extrémums est la suivante :

PROFooNVKXooXtKkuz

### Proposition 17.65.

Soit  $f: A \subset \mathbb{R} \rightarrow \mathbb{R}$  et  $a \in \text{Int}(A)$ . Supposons que  $f$  est dérivable en  $a$ . Si  $a$  est un extrémum local<sup>36</sup>, alors  $f'(a) = 0$ .

La réciproque n'est pas vraie, comme le montre l'exemple de la fonction  $x \mapsto x^3$  en  $x = 0$  : sa dérivée est nulle et pourtant  $x = 0$  n'est ni un maximum ni un minimum local.

Cette proposition ne sert donc qu'à sélectionner des *candidats* extrémum. Afin de savoir si ces candidats sont des extrémums, il y a la proposition suivante.

PROPooCXHPooSIRsEJ

### Proposition 17.66.

Soit  $f: I \subset \mathbb{R} \rightarrow \mathbb{R}$ , une fonction de classe  $C^k$  au voisinage d'un point  $a \in \text{Int } I$ . Supposons que

$$f'(a) = f''(a) = \dots = f^{(k-1)}(a) = 0, \quad (17.273)$$

et que

$$f^{(k)}(a) \neq 0. \quad (17.274)$$

Dans ce cas,

- (1) Si  $k$  est pair, alors  $a$  est un point d'extrémum local de  $f$ , c'est un minimum si  $f^{(k)}(a) > 0$ , et un maximum si  $f^{(k)}(a) < 0$ ,
- (2) Si  $k$  est impair, alors  $a$  n'est pas un extrémum local de  $f$ .

Note : jusqu'à présent nous n'avons rien dit des extrémums *globaux* de  $f$ . Il n'y a pas grand chose à en dire. Si un point d'extrémum global est situé dans l'intérieur du domaine de  $f$ , alors il sera extrémum local (a fortiori). Ou alors, le maximum global peut être sur le bord du domaine. C'est ce qui arrive à des fonctions strictement croissantes sur un domaine compact.

36. Définition 17.63.

Une seule certitude : si une fonction est continue sur un compact, elle possède une minimum et un maximum global par le théorème 9.56.

Soit une fonction  $f: I \rightarrow \mathbb{R}$ , et soit  $a \in I$ . Si  $f'(a) > 0$ , alors la tangente au graphe de  $f$  au point  $(a, f(a))$  sera une droite croissante (coefficient directeur positif). Cela ne veut pas spécialement dire que la fonction elle-même sera croissante, mais en tout cas, cela est un bon indice.

**Exemple 17.67.**

Si  $f(x) = x^2$ , il est connu que  $f'(x) = 2x$ . Nous avons donc que  $f'$  est positive si  $x \geq 0$  et  $f'$  est négative si  $x < 0$ . Cela correspond bien au fait que  $x^2$  est décroissante sur  $] -\infty, 0[$  et croissante sur  $]0, \infty[$ .  $\triangle$

Sur la figure 17.2, nous avons dessiné la fonction  $f(x) = x \cos(x)$  et sa dérivée. Nous voyons que partout où la dérivée est négative, la fonction est décroissante tandis que, inversement, partout où la dérivée est positive, la fonction est croissante.

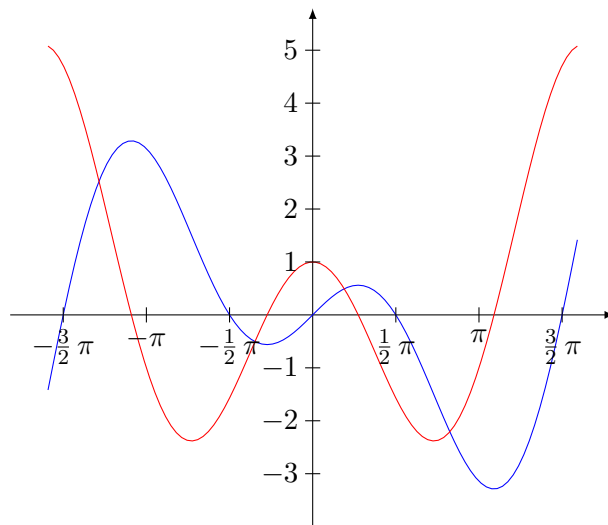


FIGURE 17.2: La fonction  $f(x) = x \cos(x)$  en bleu et sa dérivée en rouge.

Les extrémums de la fonction  $f$  sont donc placés là où  $f'$  change de signe. En effet si  $f'(x) < 0$  pour  $x < a$  et  $f'(x) > 0$  pour  $x > a$ , la fonction est décroissante jusqu'à  $a$  et est ensuite croissante. Cela signifie que la fonction connaît un creux en  $a$ . Le point  $a$  est donc un minimum de la fonction.

Attention cependant. Le fait que  $f'(a) = 0$  ne signifie pas automatiquement que  $f$  a un maximum ou un minimum en  $a$ . Nous avons par exemple tracé sur la figure 17.3 les fonctions  $x^3$  et sa dérivée. Il est à noter que, conformément à ce que l'on pense, certes la dérivée s'annule en  $x = 0$ , mais elle ne change pas de signe.

### 17.11.2 Extrema libre

**Définition 17.68.**

Un point  $a$  à l'intérieur du domaine d'une fonction  $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est un **point critique** de  $f$  lorsque  $df(a) = 0$ .

Ces points sont analogues aux points où la dérivée d'une fonction sur  $\mathbb{R}$  s'annule. Les points critiques de  $f$  sont donc les candidats à être des points d'extrémum.

Dans le cas d'une fonction de deux variables, la proposition 12.375 nous permet de voir  $(d^2f)_a$  comme étant la matrice

$$d^2f(a) = \begin{pmatrix} \frac{d^2f}{dx^2}(a) & \frac{d^2f}{dx dy}(a) \\ \frac{d^2f}{dy dx}(a) & \frac{d^2f}{dy^2}(a) \end{pmatrix}. \quad (17.275)$$

Dans le cas d'une fonction  $C^2$ , cette matrice est symétrique.

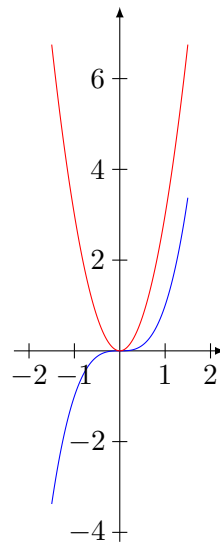


FIGURE 17.3: La dérivée de  $x^3$  s'annule en  $x = 0$ , mais ce n'est ni un minimum ni un maximum.

### Proposition 17.69 ([574]).

Soit un ouvert  $\Omega$  de  $\mathbb{R}^n$  et  $a \in \Omega$ . Soit une fonction  $f: \Omega \rightarrow \mathbb{R}$  différentiable en  $a$ . Si  $a$  est un extrémum local de  $f$ , alors  $a$  est un point critique de  $f$ .

*Démonstration.* Nous supposons que  $a$  est un maximum local (ce sera la même chose si  $a$  est un minimum). Soit  $r > 0$  tel que  $f(x) \leq f(a)$  pour tout  $x \in B(a, r)$  (et tel que cette boule reste dans  $\Omega$ ). Soit  $u \in \mathbb{R}^n$  assez petit pour que  $a \pm u \in B(a, r)$  de sorte que la définition suivante ait un sens :

$$\begin{aligned} g: [-1, 1] &\rightarrow \mathbb{R} \\ t &\mapsto f(a + tu) \end{aligned} \tag{17.276}$$

Cette fonction est différentiable en  $t = 0$  (composée de fonctions différentiables, théorème 10.303) et a un maximum local en  $t = 0$ . Donc  $g'(0) = 0$  par la proposition 17.65. Donc

$$0 = \frac{d}{dt} \left[ f(a + tu) \right]_{t=0} = df_a(u). \tag{17.277}$$

□

### 17.11.3 Extremums et Hessienne

#### Proposition 17.70 ([1, 575, 576]).

Soit un ouvert  $\Omega$  de  $\mathbb{R}^n$  et une fonction  $f: \Omega \rightarrow \mathbb{R}$  deux fois différentiable ainsi que  $a \in \Omega$ .

- (1) Si  $a$  est un point critique de  $f$  et si il existe  $r$  tel que  $(d^2 f_x)$  est semi-définie positive pour tout  $x \in B(a, r)$  alors  $f$  possède un minimum local en  $a$ .
- (2) Si  $a$  est un point critique<sup>37</sup> de  $f$ , et si  $d^2 f_a$  est strictement définie positive<sup>38</sup>, alors  $a$  est un minimum local strict de  $f$ ,
- (3) Si  $a$  est un minimum local, alors  $(d^2 f)_a$  est semi-définie positive.

*Démonstration.* Nous subdivisons la preuve.

37. Définition 17.68.

38. La fonction  $f$  est de classe  $C^2$ , donc les dérivées croisées sont égales et  $d^2 f$  est symétrique. La définition 8.229 s'applique donc.

(1) **(1)**

Soit  $h$  tel que  $a + h \in B(a, r)$ . Nous allons montrer que  $f(a) \leq f(a + h)$ ; cela montrera que  $x = a$  est un minimum local. Pour cela nous utilisons un développement de Taylor<sup>39</sup> : il existe  $c \in ]a, a + h[$  tel que

$$f(a + h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_c(h, h) \geq f(a) \quad (17.278)$$

parce que, par hypothèse,  $(d^2f)_c$  est définie positive, et parce que  $df_a = 0$ .

(2) **(2)**

La forme bilinéaire  $d^2f_a$  est strictement définie positive, donc il existe  $\alpha > 0$  tel que

$$d^2f_a(h, h) > \alpha \|h\|^2 \quad (17.279)$$

pour tout  $h$ . Nous utilisons encore Taylor : il existe une fonction  $\epsilon$  telle que  $\lim_{h \rightarrow 0} \epsilon(h) = 0$  et

$$f(a + h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h, h) + \|h\|^2 \epsilon(h). \quad (17.280)$$

En tenant compte du fait que  $df_a = 0$ ,

$$f(a + h) > f(a) + \|h\|^2 \left( \frac{1}{2} \alpha + \epsilon(h) \right). \quad (17.281)$$

La limite de  $\epsilon$  nous dit qu'il existe  $r > 0$  tel que  $\|\epsilon(h)\| < \frac{1}{2} \alpha$  pour tout  $h \in B(0, r)$ . Pour ces valeurs de  $h$  nous avons

$$f(a + h) > f(a). \quad (17.282)$$

Donc  $a$  est un minimum local strict de  $f$ .

(3) **(3)** Si  $a$  est un minimum local, nous savons déjà  $df_a = 0$  par la proposition 17.69. Nous écrivons le développement de Taylor de  $f$  à l'ordre 2 de la proposition 12.505 :

$$f(a + h) = f(a) + df_a(h) + \frac{1}{2}(d^2f)_a(h, h) + \|h\|^2 \alpha(\|h\|). \quad (17.283)$$

En prenant  $h$  assez petit pour que  $a + h$  ne sorte pas de la boule dans laquelle  $a$  est un minimum, nous avons  $f(a + h) - f(a) > 0$ . Donc

$$\frac{1}{2}(d^2f)_a(h, h) + \|h\|^2 \alpha(\|h\|) > 0 \quad (17.284)$$

Nous divisons cela par  $\|h\|^2$  et notons  $e_h = h/\|h\|$  :

$$\frac{1}{2}(d^2f)_a(e_h, e_h) + \alpha(\|h\|) > 0. \quad (17.285)$$

À la limite  $h \rightarrow 0$ , le premier terme est constant tandis que le deuxième tend vers zéro. À la limite,

$$(d^2f)_a(e_h, e_h) \geq 0. \quad (17.286)$$

La caractérisation du lemme 8.233(2) nous dit alors que  $(d^2f)_a$  est semi-définie positive.

□

La partie (3) est tout à fait comparable au fait bien connu que, pour une fonction  $f: \mathbb{R} \rightarrow \mathbb{R}$ , si le point  $a$  est minimum local, alors  $f'(a) = 0$  et  $f''(a) > 0$ .

Notons que le point (3) ne parle pas de minimum strict, et donc pas de matrice *strictement* définie positive.

---

39. Proposition 12.508.

**Exemple 17.71** (Proposition 17.70(2) sans point critique).

L'hypothèse de point critique pour l'utilisation de la stricte définition positive de  $d^2f_a$  est nécessaire. Soit en effet la fonction

$$f(x) = x^2 + x. \quad (17.287)$$

Elle vérifie  $f''(0) = 2$ , de telle sorte que sa différentielle seconde en zéro soit strictement définie positive. Le point  $x = 0$  n'est cependant même pas un minimum local. Entre autres parce que  $f'(0) = 1 \neq 0$ .  $\triangle$

La méthode pour chercher les extrémums de  $f$  est donc de suivre les points suivants :

- (1) Trouver les candidats extrémums en résolvant  $\nabla f = (0, 0)$ ,
- (2) écrire  $d^2f(a)$  pour chacun des candidats
- (3) calculer les valeurs propres de  $d^2f(a)$ , déterminer si la matrice est définie positive ou négative,
- (4) conclure.

Une conséquence de la proposition 8.232(3)<sup>40</sup> est que si  $\det M < 0$ , alors le point  $a$  n'est pas un extrémum dans le cas où  $M = d^2f(a)$  par le point (3) de la proposition 17.70.

**Exemple 17.72.**

Soit la fonction  $f(x, y) = x^4 + y^4 - 4xy$ . C'est une fonction différentiable sans problème. D'abord sa différentielle est

$$df = (4x^3 - 4y; 4y^3 - 4x), \quad (17.288)$$

et la matrice des dérivées secondes est

$$M = d^2f(x, y) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}. \quad (17.289)$$

Nous avons  $df = 0$  pour les trois points  $(0, 0)$ ,  $(1, 1)$  et  $(1, -1)$ .

Pour le point  $(0, 0)$  nous avons

$$M = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix}, \quad (17.290)$$

dont les valeurs propres sont 4 et  $-4$ . Elle n'est donc ni définie ni semi-définie positive ou négative. La proposition 17.70(2) conclut donc  $(0, 0)$  n'est pas un extrémum local.

Au contraire pour les points  $(1, 1)$  et  $(-1, -1)$  nous avons

$$M = \begin{pmatrix} 12 & -4 \\ -4 & 12 \end{pmatrix}, \quad (17.291)$$

dont les valeurs propres sont 16 et 8. La matrice  $d^2f$  y est donc définie positive. Ces deux points sont donc extrémums locaux.  $\triangle$

#### 17.11.4 Un peu de recettes de cuisine

- (1) Rechercher les points critiques, c'est-à-dire les  $(x, y)$  tels que

$$\begin{cases} \frac{\partial f}{\partial x}(x, y) = 0 \\ \frac{\partial f}{\partial y}(x, y) = 0 \end{cases}$$

En effet, si  $(x_0, y_0)$  est un extrémum local de  $f$ , alors  $\frac{\partial f}{\partial x}(x_0, y_0) = 0 = \frac{\partial f}{\partial y}(x_0, y_0)$ .

- (2) Déterminer la nature des points critiques : « test » des dérivées secondes :

$$\text{On pose } H(x_0, y_0) = \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \frac{\partial^2 f}{\partial y^2}(x_0, y_0) - \left( \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \right)^2$$

(2a) Si  $H(x_0, y_0) > 0$  et  $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0 \implies (x_0, y_0)$  est un minimum local de  $f$ .

(2b) Si  $H(x_0, y_0) > 0$  et  $\frac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0 \implies (x_0, y_0)$  est un maximum local de  $f$ .

(2c) Si  $H(x_0, y_0) < 0 \implies f$  a un point de selle en  $(x_0, y_0)$ .

(2d) Si  $H(x_0, y_0) = 0 \implies$ , on ne peut rien conclure.

---

40. La matrice  $d^2f(a)$  est toujours symétrique quand  $f$  est de classe  $C^2$ .

### 17.11.5 Extrema liés

Soit  $f$ , une fonction sur  $\mathbb{R}^n$ , et  $M \subset \mathbb{R}^n$  une variété de dimension  $m$ . Nous voulons savoir quelles sont les plus grandes et plus petites valeurs atteintes par  $f$  sur  $M$ .

Pour ce faire, nous avons un théorème qui permet de trouver des extrémums *locaux* de  $f$  sur la variété. Pour rappel,  $a \in M$  est un **extrémum local de  $f$  relativement** à l'ensemble  $M$  si il existe une boule  $B(a, \epsilon)$  telle que  $f(a) \leq f(x)$  pour tout  $x \in B(a, \epsilon) \cap M$ .

ThoRGJosS

**Théorème 17.73** (Extremum lié [555]).

Soit  $A$ , un ouvert de  $\mathbb{R}^n$  et

- (1) une fonction (celle à minimiser)  $f \in C^1(A, \mathbb{R})$ ,
- (2) des fonctions (les contraintes)  $G_1, \dots, G_r \in C^1(A, \mathbb{R})$ ,
- (3)  $M = \{x \in A \text{ tel que } G_i(x) = 0 \forall i\}$ ,
- (4) un extrémum local  $a \in M$  de  $f$  relativement à  $M$ .

Supposons que les gradients  $\nabla G_1(a), \dots, \nabla G_r(a)$  soient linéairement indépendants. Alors  $a = (x_1, \dots, x_n)$  est une solution de  $\nabla L(a) = 0$  où

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_r) = f(x_1, \dots, x_n) + \sum_{i=1}^r \lambda_i G_i(x_1, \dots, x_n). \quad (17.292)$$

Autrement dit, si  $a$  est un extrémum lié, alors  $\nabla f(a)$  est une combinaison des  $\nabla G_i(a)$ , ou encore il existe des  $\lambda_i$  tels que

$$df(a) = \sum_i \lambda_i dG_i(a). \quad (17.293)$$

EqRDsSXvZ

La fonction  $L$  est le **lagrangien** du problème et les variables  $\lambda_i$  sont les **multiplicateurs de Lagrange**.

*Démonstration.* Si  $r = n$  alors les vecteurs linéairement indépendants  $\nabla G_i(a)$  forment une base de  $\mathbb{R}^n$  et donc évidemment les  $\lambda_i$  existent. Nous supposons donc maintenant que  $r < n$ . Nous notons  $(z_i)_{i=1 \dots n}$  les coordonnées sur  $\mathbb{R}^n$ .

La matrice

$$\begin{pmatrix} \frac{\partial G_1}{\partial z_1}(a) & \dots & \frac{\partial G_1}{\partial z_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial G_r}{\partial z_1}(a) & \dots & \frac{\partial G_r}{\partial z_n}(a) \end{pmatrix} \quad (17.294)$$

est de rang  $r$  parce que les lignes sont par hypothèses linéairement indépendantes. Nous nommons  $(y_i)_{i=1, \dots, r}$  un choix de  $r$  parmi les  $(z_i)$  tels que

$$\det \begin{pmatrix} \frac{\partial G_1}{\partial y_1} & \dots & \frac{\partial G_1}{\partial y_r} \\ \vdots & \ddots & \vdots \\ \frac{\partial G_r}{\partial y_1} & \dots & \frac{\partial G_r}{\partial y_r} \end{pmatrix} \neq 0. \quad (17.295)$$

Nous identifions  $\mathbb{R}^n$  à  $\mathbb{R}^s \times \mathbb{R}^r$  dans lequel  $\mathbb{R}^r$  est la partie générée par les  $(y_i)_{i=1, \dots, r}$ . Les coordonnées sur  $\mathbb{R}^s$  seront nommées  $(x_j)_{j=1, \dots, s}$ , de telle sorte que les coordonnées sur  $\mathbb{R}^n$  seront  $x_1, \dots, x_s, y_1, \dots, y_r$ . Dans ces coordonnées, nous nommons  $a = (\alpha, \beta)$  avec  $\alpha \in \mathbb{R}^s$  et  $\beta \in \mathbb{R}^r$ .

Si nous notons  $G = (G_1, \dots, G_r)$ , le théorème de la fonction implicite (théorème 17.48) nous dit qu'il existe un voisinage  $U'$  de  $\alpha \in \mathbb{R}^s$ , un voisinage  $V'$  de  $\beta \in \mathbb{R}^r$  et une fonction  $\varphi: U' \rightarrow V'$  de classe  $C^1$  telle que si  $(x, y) \in U' \times V'$ , alors

$$G(x, y) = 0 \quad (17.296)$$

si et seulement si  $y = \varphi(x)$ . Nous posons maintenant

$$\psi(x) = (x, \varphi(x)) \quad (17.297a)$$

$$h(x) = f(\psi(x)). \quad (17.297b)$$



Nous avons  $\psi(\alpha) = a$  et  $\psi(x) \in M$  pour tout  $x \in U'$ . La fonction  $h$  a donc un extrémum local en  $\alpha$  et donc les dérivées partielles de  $h$  y sont nulles. Cela signifie que

$$0 = \frac{\partial h}{\partial x_i}(\alpha) = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial x_i} + \sum_{k=1}^r \frac{\partial f}{\partial y_k} \frac{\partial \varphi_k}{\partial x_i}, \quad (17.298)$$

c'est-à-dire

$$\frac{\partial f}{\partial x_i}(\alpha) + \sum_{k=1}^r \frac{\partial f}{\partial y_k}(a) \frac{\partial \varphi_k}{\partial x_i}(\alpha) = 0 \quad (17.299)$$

pour tout  $i = 1, \dots, s$ . D'autre part pour tout  $k$ , la fonction  $l_k(x) = G_k(x, \varphi(x))$  est constante et vaut zéro ; ses dérivées partielles sont donc nulles :

$$\frac{\partial l}{\partial x_i}(\alpha) = \frac{\partial G_k}{\partial x_i}(\alpha) + \sum_{k=1}^r \frac{\partial G_k}{\partial y_k}(a) \frac{\partial \varphi_k}{\partial x_i}(\alpha) = 0 \quad (17.300)$$

pour tout  $i = 1, \dots, s$  et  $k = 1, \dots, r$ .

Les  $s$  premières colonnes de la matrice

$$\begin{pmatrix} \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_s} & \frac{\partial f}{\partial y_1} & \dots & \frac{\partial f}{\partial y_r} \\ \frac{\partial G_1}{\partial x_1} & \dots & \frac{\partial G_1}{\partial x_s} & \frac{\partial G_1}{\partial y_1} & \dots & \frac{\partial G_1}{\partial y_r} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial G_r}{\partial x_1} & \dots & \frac{\partial G_r}{\partial x_s} & \frac{\partial G_r}{\partial y_1} & \dots & \frac{\partial G_r}{\partial y_r} \end{pmatrix} \quad (17.301)$$

s'expriment en termes des  $r$  dernières. La matrice est donc au maximum de rang  $r$ . Notons que la première ligne est  $\nabla f$  et les  $r$  suivantes sont les  $\nabla G_i$ . Vu que ces lignes sont des vecteurs liés, il existe  $\mu_0, \dots, \mu_r$  tels que

$$\mu_0 \nabla f + \sum_{i=1}^r \mu_i \nabla G_i = 0. \quad (17.302)$$

Par hypothèse les  $\nabla G_i$  sont linéairement indépendants, ce qui nous dit que  $\mu_0 \neq 0$ . Donc nous avons ce qu'il nous faut :

$$\nabla f(a) = \sum_i \frac{\mu_i}{\mu_0} \nabla G_i(a). \quad (17.303)$$

Notons qu'au vu de l'expression (17.293), le fait que les formes  $\{dG_i(a)\}_{1 \leq i \leq r}$  forment une partie libre dans  $(\mathbb{R}^n)^*$  implique que les  $\lambda_i$  sont uniques.  $\square$

La proposition suivante est la même que 17.73.

PropfPUxh

### Proposition 17.74.

Soit  $U$ , un ouvert de  $\mathbb{R}^n$  et des fonctions de classe  $C^1$   $f, g_1, \dots, g_r : U \rightarrow \mathbb{R}$ . Nous considérons

$$\Gamma = \{x \in U \text{ tel que } g_1(x) = \dots = g_r(x) = 0\}. \quad (17.304)$$

Soit  $a$  un extrémum de  $f|_{\Gamma}$ . Supposons que les formes  $dg_1, \dots, dg_r$  soient linéairement indépendantes en  $a$ . Alors il existe  $\lambda_1, \dots, \lambda_r$  dans  $\mathbb{R}$  tel que

$$df_a = \sum_{i=1}^r \lambda_i (dg_i)_a. \quad (17.305)$$

En pratique les candidats extrémums locaux sont tous les points où les gradients ne sont pas linéairement indépendants, plus tous les points donnés par l'équation  $\nabla L = 0$ . Parmi ces candidats, il faut trouver lesquels sont maximums ou minima, locaux ou globaux.

L'existence d'extrémums locaux se prouve généralement en invoquant de la compacité, et en invoquant le lemme suivant qui permet de réduire le problème à un compact.

LemmeMinSCimpliqueS

**Lemme 17.75.**

Soit  $S$ , une partie de  $\mathbb{R}^n$  et  $C$ , un ouvert de  $\mathbb{R}^n$ . Si  $a \in \text{Int } S$  est un minimum local relatif à  $S \cap C$ , alors il est un minimum local par rapport à  $S$ .

*Démonstration.* Vu que  $a$  est un minimum local relatif à  $S \cap C$ , il existe un  $\epsilon_1 > 0$  tel que pour tout  $x \in B(a, \epsilon_1) \cap S \cap C$ , nous avons  $f(x) \geq f(a)$ .

Mais étant donné que  $C$  est ouvert, et que  $a \in C$ , il existe un  $\epsilon_2$  tel que  $B(a, \epsilon_2) \subset C$ . En prenant  $\epsilon = \min\{\epsilon_1, \epsilon_2\}$ , nous trouvons que  $f(x) \geq f(a)$  pour tout  $x \in B(a, \epsilon) \cap (S \cap C) = B(a, \epsilon) \cap S$ .  $\square$

**17.12 Fonctions convexes**

SECooVZWwQXRXQa

**Définition 17.76** ([577]).

Une fonction  $f$  d'un intervalle  $I$  de  $\mathbb{R}$  vers  $\mathbb{R}$  est dite **convexe** lorsque, pour tous  $x_1$  et  $x_2$  de  $I$  et tout  $\lambda$  dans  $[0, 1]$  nous avons

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \quad (17.306)$$

EQooYNAPooFePOZy

Si pour tout  $\lambda \in ]0, 1[$  et pour tout  $x \neq y$  dans  $I$  nous avons

$$f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2) \quad (17.307)$$

EQooEGBSoocvopB

alors nous disons que la fonction  $f$  est **strictement convexe** sur  $I$ .

Une fonction est **concave** si son opposée est convexe.

**17.77** ([577]).

Les différents résultats pour les fonctions convexes s'adaptent généralement sans mal aux fonctions strictement convexes. Une nuance cependant : de même que les fonctions dérivables convexes sont celles qui ont une dérivée croissante, les fonctions dérivables strictement convexes sont celles qui ont une dérivée strictement croissante (proposition 17.83). En revanche, il ne faudrait pas croire que la dérivée seconde d'une fonction dérivable strictement convexe est nécessairement une fonction à valeurs strictement positives (voir théorème 17.84) : la dérivée d'une fonction strictement croissante peut s'annuler occasionnellement, ou plus exactement peut s'annuler sur un ensemble de points d'intérieur vide. Penser à  $x \mapsto x^4$  pour un exemple de fonction strictement convexe dont la dérivée seconde s'annule.

**17.78.**

Demander un  $\lambda \in [0, 1]$  est équivalent à demander des  $\alpha, \beta \geq 0$  tels que  $\alpha + \beta = 1$ . En effet, si  $\alpha + \beta = 1$ , alors il suffit de se dire que  $\alpha$  est le  $\lambda$  et que  $\beta = 1 - \alpha$ .

LEMooXUDSoocBUMnaZ

**Lemme 17.79.**

soit une fonction strictement convexe  $f$ . Soient  $\alpha, \beta > 0$  tels que  $\alpha + \beta = 1$ . Nous avons

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v) \quad (17.308)$$

EQooDKKUooSHLfwy

si et seulement si  $u = v$ .

*Démonstration.* Tant que  $u \neq v$ , la définition (17.307) est une inégalité stricte. Donc si (17.308) est valide, alors  $u = v$ .

Dans le sens inverse, si  $u = v$ , alors  $\alpha u + \beta v = u$ , et  $\alpha f(u) + \beta f(v) = f(u)$ .  $\square$

**17.12.1 Inégalité des pentes**

Dans l'étude des fonctions convexes nous allons souvent utiliser la fonction **taux d'accroissement** qui est, pour  $\alpha$  dans le domaine de convexité de  $f$  définie par

$$\tau_\alpha: I \setminus \{\alpha\} \rightarrow \mathbb{R}$$

$$x \mapsto \frac{f(x) - f(\alpha)}{x - \alpha}.$$

EaRYBazWd

(17.309)

**Proposition 17.80** (Inégalité des pentes[578]).

Soit  $f$  une fonction convexe sur un intervalle  $I \subset \mathbb{R}$ . Alors pour tout  $a < b < c$  dans  $I$  nous avons<sup>41</sup>

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}. \quad (17.310)$$

En d'autres termes,

$$\tau_a(b) \leq \tau_a(c) \leq \tau_b(c), \quad (17.311)$$

c'est-à-dire que  $\tau$  est croissante en ses deux arguments.

*Démonstration.* D'abord les inégalités  $a < b < c$  impliquent  $0 < b - a < c - a$  et donc

$$\lambda = \frac{b - a}{c - a} < 1. \quad (17.312)$$

L'astuce est de remarquer que  $(1 - \lambda)a + \lambda c = b$ . Donc  $\lambda$  a toutes les bonnes propriétés pour être utilisé dans la définition de la convexité :

$$f((1 - \lambda)a + \lambda c) \leq \lambda f(c) + (1 - \lambda)f(a), \quad (17.313)$$

c'est-à-dire

$$f(b) - f(a) \leq \lambda(f(c) - f(a)) \quad (17.314)$$

ou encore, en remplaçant  $\lambda$  par sa valeur :

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a}. \quad (17.315)$$

Cela fait déjà une des inégalités à savoir.

D'autre part en partant de  $-a > -b < c$  nous posons

$$0 < \lambda = \frac{c - b}{c - a}. \quad (17.316)$$

Nous avons à nouveau  $b = (1 - \lambda)c + \lambda a$  et nous pouvons obtenir la seconde inégalité

$$\frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}. \quad (17.317)$$

□

Géométriquement, l'inégalité des pentes se comprend facilement : le coefficient angulaire de la corde du graphe augmente. Donc si  $x < y < z$ , le coefficient moyen entre  $x$  et  $y$  est plus petit que celui entre  $x$  et  $z$  qui est plus petit que celui entre  $y$  et  $z$ .

Donc si le coefficient angulaire moyen entre  $a$  et  $b + u$  vaut celui entre  $a$  et  $b$ , ce coefficient ne peut qu'être constant entre  $a$  et  $b$  : sinon il serait plus grand entre  $b$  et  $b + u$  et la moyenne sur  $a \rightarrow b + u$  serait plus grande que sa moyenne sur  $a \rightarrow b$ . Mais avoir un coefficient angulaire constant signifie être une droite.

En résumé, si une fonction est convexe et non strictement convexe, alors son graphe est une droite. C'est en gros cela que la proposition 17.88 clarifiera.

---

41. Les inégalités sont strictes si la fonction  $f$  est strictement convexe.

### 17.12.2 Convexité et régularité

LemKLTsHIQ

**Lemme 17.81** ([577]).

Une fonction convexe sur un ouvert de  $\mathbb{R}$

- (1)  $y$  admet des dérivées à gauche et à droite en chaque point,
- (2)  $y$  est continue.

*Démonstration.* Si  $A$  est l'ouvert sur lequel la fonction  $f$  est convexe, pour chaque point  $x$  de  $A$ , il existe une boule  $B(x, r) \subset A$ . Pour  $0 < r' < r$  nous avons aussi  $\overline{B(0, r')} \subset A$ . Bref, nous pouvons supposer que  $f$  est définie et convexe sur un intervalle  $I = [a, b]$ .

Soit  $\alpha \in I$ . Nous allons prouver que  $f$  est continue en  $\alpha$ . Nous considérons  $\tau_\alpha$  le taux d'accroissement défini par (17.309); c'est une fonction croissante comme précisé dans l'inégalité des trois pentes 17.80 et de plus  $\tau_\alpha(x)$  est bornée supérieurement par  $\tau_\alpha(b)$  pour  $x < \alpha$  et inférieurement par  $\tau_\alpha(a)$  pour  $x > \alpha$ . Les limites existent donc et sont finies par la proposition 12.58. Autrement dit les limites

$$\lim_{x \rightarrow \alpha^+} \frac{f(x) - f(\alpha)}{x - \alpha} = \lim_{x \rightarrow \alpha^+} \tau_\alpha(x) = \inf_{t > \alpha} \tau_\alpha(t) \quad (17.318a)$$

$$\lim_{x \rightarrow \alpha^-} \frac{f(x) - f(\alpha)}{x - \alpha} = \lim_{x \rightarrow \alpha^-} \tau_\alpha(x) = \sup_{t < \alpha} \tau_\alpha(t). \quad (17.318b)$$

existent et sont finies, c'est-à-dire que la fonction  $f$  admet une dérivée à gauche et à droite.

Pour tout  $x$  nous avons les inégalités

$$\tau_\alpha(a) \leq \frac{f(x) - f(\alpha)}{x - \alpha} \leq \tau_\alpha(b). \quad (17.319)$$

En posant  $k = \max\{\tau_\alpha(a), \tau_\alpha(b)\}$  nous avons

$$|f(x) - f(\alpha)| \leq k|x - \alpha|. \quad (17.320)$$

La fonction est donc Lipschitzienne et par conséquent continue par la proposition 12.344.  $\square$

**Remarque 17.82.**

Les dérivées à gauche et à droite ne sont à priori pas égales. Penser par exemple à une fonction affine par morceaux dont les pentes augmentent à chaque morceau.

### 17.12.3 Dérivées d'une fonction convexe

PropYKwTDPX

**Proposition 17.83** ([579, 580, 1]).

Une fonction dérivable sur un intervalle  $I$  de  $\mathbb{R}$

ITEMooUTSAooJvhZNm

- (1) est convexe si et seulement si sa dérivée est croissante sur  $I$ .

ITEMooLLSIooFwkxtV

- (2) est strictement convexe si et seulement si sa dérivée est strictement croissante sur  $I$

*Démonstration.* Pour la preuve de (1) et (2), nous allons démontrer les énoncés « non stricts » et indiquer ce qu'il faut changer pour obtenir les énoncés « stricts ».

- (1) **Sens direct** Nous supposons que  $f$  est convexe. Soient  $a < b$  dans  $I$  et  $x \in ]a, b[$ . D'après l'inégalité des pentes 17.80,

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(a)}{b - a} \leq \frac{f(b) - f(x)}{b - x}. \quad (17.321) \quad \text{EqATDLooIcqdDI}$$

En faisant la limite  $x \rightarrow a$  nous avons

$$f'(a) \leq \frac{f(b) - f(a)}{b - a} \quad (17.322)$$

et la limite  $x \rightarrow b$  donne

$$\frac{f(b) - f(a)}{b - a} \leq f'(b). \quad (17.323)$$

Ici les inégalités sont à priori non-strictes, même si  $f$  est strictement convexe : même avec des inégalités strictes dans (17.321), le passage à la limite rend l'inégalité non stricte. Quoi qu'il en soit nous avons

$$f'(a) \leq f'(b). \quad \text{EqQGVMOoBpuvNr} \quad (17.324)$$

- (2) **Sens direct : strict** Nous savons déjà que  $f'$  est croissante. Si (17.324) était une égalité, alors  $f'$  serait constante sur  $]a, b[$  parce qu'en prenant  $c$  entre  $a$  et  $b$  nous aurions  $f'(a) \leq f'(c) \leq f'(b)$  avec  $f'(a) = f'(b)$ . Donc  $f'(a) = f'(c)$ . Avoir  $f'$  constante sur un intervalle est contraire à la stricte convexité.

(3) **Sens réciproque**

Nous supposons que  $f'$  est croissante et nous considérons  $a < b$  dans  $I$  ainsi que  $\lambda \in [0, 1]$ . Nous posons  $x = \lambda a + (1 - \lambda)b$ , et nous savons que  $a \leq x \leq b$ . Le théorème des accroissements finis 12.216 donne  $c_1 \in ]a, x[$  et  $c_2 \in ]x, b[$  tels que

$$f'(c_1) = \frac{f(x) - f(a)}{x - a} \quad (17.325)$$

et

$$f'(c_2) = \frac{f(b) - f(x)}{b - x}. \quad (17.326)$$

Et en plus  $c_1 < c_2$ . Vu que  $f'$  est croissante nous avons  $f'(c_1) \leq f'(c_2)$  et donc

$$\frac{f(x) - f(a)}{x - a} \leq \frac{f(b) - f(x)}{b - x}. \quad \text{EqSAOCooWAwCLO} \quad (17.327)$$

En remplaçant  $x$  par sa valeur en termes de  $\lambda$ ,  $a$  et  $b$  nous avons  $x - a = (1 - \lambda)(b - a)$  et  $b - x = \lambda(b - a)$ , et l'inégalité (17.327) nous donne

$$f(x) \leq \lambda f(a) + (1 - \lambda)f(b). \quad (17.328)$$

- (4) **Sens réciproque : strict** Si  $f'$  est strictement croissante, nous avons  $f'(c_1) < f'(c_2)$  et les inégalité suivantes sont strictes, ce qui donne

$$f(x) < \lambda f(a) + (1 - \lambda)f(b). \quad (17.329)$$

□

ThoGXjKeYb

**Théorème 17.84** ([579]).

Soit une fonction  $f$  de classe  $C^2$ .

- (1) Est convexe si et seulement si  $f''$  est positive. ITEMooIUTQooTkRMoyBP  
 (2) Si  $f''$  est strictement positive, elle est strictement convexe. ITEMooXUOMooYIoOtv

*Démonstration.* En deux parties.

- (1) **Pour (1)** La fonction est  $C^2$ , donc  $f''$  est positive si et seulement si  $f'$  est croissante (proposition 12.209) alors que la proposition 17.83 nous jure que  $f$  sera convexe si et seulement si  $f'$  est croissante.  
 (2) **Pour (2)** Si  $f''$  est strictement positive,  $f'$  sera strictement croissante et donc  $f$  strictement convexe (proposition 17.83).

□

REMooVRPQooIybxmp

**Remarque 17.85.**

Une fonction peut être strictement convexe sans que sa dérivée seconde ne soit toujours strictement positive. En exemple :  $x \mapsto x^4$  est strictement convexe alors que sa dérivée seconde s'annule en zéro.

ExPDRooZCtkOz

**Exemple 17.86.**

Quelques exemples utilisant le théorème 17.84

- (1) La fonction  $x \mapsto x^2$  est convexe parce que sa dérivée seconde est la constante (positive) 2.
- (2) La fonction  $x \mapsto \frac{1}{x}$  est convexe sur  $\mathbb{R}^+ \setminus \{0\}$  (sa dérivée seconde est  $2x^{-3}$ ). ITEMooRXSBooDBerbx
- (3) La fonction exponentielle est strictement convexe par le théorème 17.84.
- (4) La fonction  $\ln$  est concave parce que la dérivée seconde de  $-\ln$  est  $\frac{1}{x^2}$  qui est strictement positif.

△

Nous en démontrons une en détail ; elle sera utile en analyse fonctionnelle, lors de l'étude des espaces  $L^p$ . Voir par exemple le théorème de la projection 27.155.

LEMooSXTXooZOmtKq**Lemme 17.87.**

Soient  $p > 1$  et la fonction

$$\begin{aligned} f: ]0, \infty[ &\rightarrow \mathbb{R} \\ x &\mapsto x^p \end{aligned} \tag{17.330}$$

est strictement convexe.

*Démonstration.* La proposition 14.295 nous permet de dire que la fonction  $f$  est de classe  $C^\infty$  et que la dérivée seconde est donnée par

$$f''(x) = p(p-1)x^{p-2}. \tag{17.331}$$

Cela est strictement positif pour tous les  $x$  considérés, le théorème 17.84 conclut. □

**17.12.4 Graphe d'une fonction convexe**

L'idée principale du graphe d'une fonction convexe est qu'il est toujours au dessus du graphe de ses tangentes (lorsqu'elles existent). Lorsqu'elles n'existent pas, le lemme 17.81 donne des coefficients directeurs de droites qui vont rester en dessous du graphe de la fonction.

PROPooOCOEOoEGybmS**Proposition 17.88 ([581]).**

Une fonction convexe est strictement convexe si et seulement si il n'existe aucun intervalle de longueur non nulle sur lequel elle coïncide avec une fonction affine.

*Démonstration.* Si sur l'intervalle (non réduit à un point)  $[x, y]$ , la fonction convexe  $f$  coïncide avec une fonction affine, alors  $f(t) = at + b$  et pour  $\lambda \in ]0, 1[$  nous avons

$$f(\lambda x + (1-\lambda)y) = a\lambda x + a(1-\lambda)y + b = \lambda f(x) + (1-\lambda)f(y) \tag{17.332}$$

où nous avons remplacé  $b$  par  $\lambda b + (1-\lambda)b$ . Par conséquent la fonction n'est pas strictement convexe.

Nous supposons maintenant que la fonction convexe  $f$  n'est pas strictement convexe sur l'intervalle  $I$ . Il existe  $x \neq y \in I$  et  $\lambda \in ]0, 1[$  tels que

$$f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y). \tag{17.333}$$

Nous posons  $z = \lambda x + (1-\lambda)y$  et  $u \in ]x, z[$  pour écrire des inégalités des pentes entre  $x < u < z < y$ . Plus précisément si nous notons  $a \rightarrow b$  la pente de  $a$  à  $b$ , c'est-à-dire  $a \rightarrow b = \frac{f(b)-f(a)}{b-a}$ , alors les inégalités des pentes pour  $x < u < z$  puis  $u < z < y$  donnent

$$x \rightarrow z \leq u \rightarrow z \leq z \rightarrow y. \tag{17.334}$$

EqooBMEFocMpoEzd

Voyons maintenant qu'en réalité  $z \rightarrow y = x \rightarrow z$ . En effet en remplaçant

$$f(y) = \frac{f(z) - \lambda f(x)}{1-\lambda} \tag{17.335}$$

et

$$y = \frac{\lambda x}{1 - \lambda} \quad (17.336)$$

dans l'expression  $z \rightarrow y = \frac{f(y)-f(z)}{y-z}$  nous obtenons

$$z \rightarrow y = \frac{f(y) - f(z)}{y - z} = \frac{f(z) - f(x)}{z - x} = x \rightarrow z. \quad (17.337)$$

Les inégalités (17.334) sont donc des égalités :

$$\frac{f(z) - f(x)}{z - x} = \frac{f(z) - f(u)}{z - u} = \frac{f(y) - f(z)}{y - z}. \quad (17.338)$$

Nous avons donc montré que le nombre  $a = \frac{f(z)-f(u)}{z-u}$  ne dépend pas de  $u$ . Nous avons alors

$$f(z) - f(u) = a(z - u) \quad (17.339)$$

ou encore :

$$f(u) = f(z) - a(z - u), \quad (17.340)$$

ce qui signifie que sur  $]x, z[$ , la fonction  $f$  est affine. □

PROPooQPOSooDZIUAJ

### Proposition 17.89.

Une fonction dérivable sur un intervalle  $I$  de  $\mathbb{R}$  est convexe si et seulement si son graphe est au dessus de chacune de ses tangentes.

*Démonstration.* En deux parties.

(1) **Sens direct** Soient  $x, y \in I$ . Nous voulons :

$$f(y) \geq f(x) + f'(x)(y - x). \quad (17.341)$$

Étant donné que nous aurons besoin, dans le quotient différentiel de quelque chose comme  $f(x + t) - f(x)$  nous écrivons la définition (17.306) de la convexité en inversant les rôles de  $x$  et  $y$  et en manipulant un peu :

$$f(ty + (1 - t)x) \leq tf(y) + (1 - t)f(x) \quad (17.342a)$$

$$f(x + t(y - x)) \leq tf(y) + (1 - t)f(x) \quad (17.342b)$$

$$f(x + t(y - x)) - f(x) \leq tf(y) - tf(x) \quad (17.342c)$$

Nous divisons par  $t$  :

$$\frac{f(x + t(y - x)) - f(x)}{t} \leq f(y) - f(x). \quad (17.343)$$

Le passage à la limite  $t \rightarrow 0$  donne

$$(y - x)f'(x) \leq f(y) - f(x), \quad (17.344)$$

ce qu'il fallait.

(2) **Sens inverse** Pour tout  $x, y \in I$  nous supposons avoir

$$f(y) \geq f(x) + f'(x)(y - x). \quad (17.345)$$

EQooEXXIooHXJnER

Si nous supposons  $x \neq y$  et si nous posons  $z = \lambda x + (1 - \lambda)y$  nous voulons prouver que

$$f(z) \leq \lambda f(x) + (1 - \lambda)f(y). \quad (17.346)$$

Pour cela nous écrivons l'inégalité (17.345) avec les couples  $(x, z)$  et  $(y, z)$  :

$$f(x) \geq f(z) + f'(z)(x - z) \quad (17.347a)$$

$$f(y) \geq f(z) + f'(z)(y - z) \quad (17.347b)$$

En multipliant la première par  $\lambda$  et la seconde par  $(1 - \lambda)$  et en sommant,

$$\lambda f(x) + (1 - \lambda)f(y) \geq \lambda f(z) + \lambda f'(z)(x - z) + (1 - \lambda)f(z) + (1 - \lambda)f'(z)(y - z) \quad (17.348a)$$

$$= f(z) + f'(z)(\lambda(x - z) + (1 - \lambda)(y - z)) \quad (17.348b)$$

$$= f(z). \quad (17.348c)$$

□

PropNIBooSbXIKO

**Proposition 17.90** ([1]).

Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$  une fonction convexe et  $a \in \mathbb{R}$ . Il existe une constante  $c_a \in \mathbb{R}$  telle que pour tout  $x$  nous ayons

$$f(x) - f(a) \geq c_a(x - a). \quad (17.349)$$

EqSKJooSeAekM

Autrement dit, le graphe de la fonction  $f$  est toujours au dessus de la droite d'équation

$$y = f(a) + c_a(x - a). \quad (17.350)$$

*Démonstration.* Les dérivées à gauche et à droite de  $f$  données par le lemme 17.81 sont les candidats tout désignés pour être coefficient directeur de la droite que l'on cherche. Nous allons prouver qu'en posant

$$c_a = \inf_{t > a} \tau_a(t), \quad (17.351)$$

la droite  $y = f(a) + c_a(x - a)$  répond à la question<sup>42</sup>.

Nous devons prouver que le nombre  $\Delta_x = f(x) - (f(a) + c_a(x - a))$  est positif pour tout  $x$ .

(1) Si  $x > a$

Nous divisons par  $x - a$  et nous devons prouver que  $\frac{\Delta_x}{x - a}$  est positif :

$$\frac{\Delta_x}{x - a} = \frac{f(x) - f(a)}{x - a} - c_a \quad (17.352a)$$

$$= \tau_a(x) - \inf_{t > a} \tau_a(t) \quad (17.352b)$$

$$\geq 0 \quad (17.352c)$$

parce que  $t \rightarrow \tau_a(t)$  est croissante et que  $x > a$ .

(2) Si  $x < a$

Nous divisons par  $x - a$  et nous devons prouver que  $\frac{\Delta_x}{x - a}$  est négatif :

$$\frac{\Delta_x}{x - a} = \frac{f(x) - f(a)}{x - a} - c_a \quad (17.353a)$$

$$= \tau_a(x) - \inf_{t > a} \tau_a(t) \quad (17.353b)$$

$$\leq 0 \quad (17.353c)$$

parce que  $t \rightarrow \tau_a(t)$  est croissante et que  $x < a$ .

□

PropPEJCGCH

**Proposition 17.91** ([1]).

Si  $g$  est une fonction convexe, il existe deux suites réelles  $(a_n)$  et  $(b_n)$  telles que

$$g(x) = \sup_{n \in \mathbb{N}} (a_n x + b_n). \quad (17.354)$$

42. En prenant l'autre,  $c'_a = \sup_{t < a} \tau_a(t)$ , ça fonctionne aussi. En pensant à une fonction affine par morceaux, on remarque qu'en choisissant un nombre entre les deux, nous avons plus facilement une inégalité stricte dans (17.349).



*Démonstration.* Pour  $u \in \mathbb{R}$  nous considérons  $a(u)$  et  $b(u)$  tels que la droite  $y(x) = a(u)x + b(u)$  vérifie  $y(u) = g(u)$  et  $y(x) \leq g(x)$  pour tout  $x$ . Cela est possible par la proposition 17.90. Il s'agit d'une droite coupant le graphe de  $g$  en  $x = u$  et restant en dessous. Nous considérons alors  $(u_n)$  une suite quelconque dense dans  $\mathbb{R}$  (disons les rationnels pour fixer les idées) et nous posons

$$\begin{cases} a_n = a(u_n) \\ b_n = b(u_n). \end{cases} \quad (17.355a) \quad (17.355b)$$

Si  $q \in \mathbb{Q}$  alors  $a_n x + b_n \leq g(x)$  pour tout  $n$  et  $g(q)$  est le supremum qui est atteint pour le  $n$  tel que  $u_n = q$ . Si maintenant  $x$  n'est pas dans  $\mathbb{Q}$  il faut travailler plus.

Nous prenons  $(\tilde{q}_n)$ , une sous-suite de  $(q_n)$  convergeant vers  $x$  et  $N$  suffisamment grand pour que pour tout  $n \geq N$  on ait  $|\tilde{q}_n - x| \leq \epsilon$  et  $|g(\tilde{q}_n) - g(x)| \leq \epsilon$ ; cela est possible grâce à la continuité de  $g$  (lemme 17.81). Ensuite les sous-suites  $(\tilde{a}_n)$  et  $(\tilde{b}_n)$  sont celles qui correspondent :

$$\tilde{a}_n \tilde{q}_n + \tilde{b}_n = g(\tilde{q}_n). \quad (17.356)$$

Nous considérons la majoration

$$|\tilde{a}_n x + \tilde{b}_n - g(x)| \leq |\tilde{a}_n x + \tilde{b}_n - (\tilde{a}_n \tilde{q}_n + \tilde{b}_n)| + \underbrace{|\tilde{a}_n \tilde{q}_n + \tilde{b}_n - g(\tilde{q}_n)|}_{=0} + \underbrace{|g(\tilde{q}_n) - g(x)|}_{\leq \epsilon} \quad (17.357a)$$

$$\leq |\tilde{a}_n| |x - \tilde{q}_n| + \epsilon \quad (17.357b)$$

$$= \epsilon (|\tilde{a}_n| + 1). \quad (17.357c)$$

Il nous reste à montrer que  $|\tilde{a}_n|$  est borné par un nombre ne dépendant pas de  $n$  (pour les  $n > N$ ).

Comme la droite de coefficient directeur  $\tilde{a}_n$  et passant par le point  $(\tilde{q}_n, g(\tilde{q}_n))$  reste en dessous du graphe de  $g$ , nous avons pour tout  $n$  et tout  $y \in \mathbb{R}$  l'inégalité

$$g(y) \geq \tilde{a}_n(y - \tilde{q}_n) + g(\tilde{q}_n) \in \tilde{a}_n B(y - x, \epsilon) + B(g(x), \epsilon). \quad (17.358)$$

Si  $\tilde{a}_n$  n'est pas borné vers le haut, nous prenons  $y$  tel que  $B(y - x, \epsilon)$  soit minoré par un nombre  $k$  strictement positif et nous obtenons

$$g(y) \geq k \tilde{a}_n + l \quad (17.359)$$

avec  $k$  et  $l$  indépendants de  $n$ . Cela donne  $g(y) = \infty$ . Si au contraire  $\tilde{a}_n$  n'est pas borné vers le bas, nous prenons  $y$  tel que  $B(y - x, \epsilon)$  est majoré par un nombre  $k$  strictement négatif. Nous obtenons encore  $g(y) = \infty$ .

Nous concluons que  $|\tilde{a}_n|$  est bornée. □

LemXOUooQsigHs

**Lemme 17.92** ([126]).

*L'application*

$$\begin{aligned} \phi: S^{++}(n, \mathbb{R}) &\rightarrow \mathbb{R} \\ A &\mapsto \det(A) \end{aligned} \quad (17.360)$$

est **log-convave**, c'est-à-dire que l'application  $\ln \circ \phi$  est concave<sup>43</sup>. De façon équivalente, si  $A, B \in S^{++}$  et si  $\alpha + \beta = 1$ , alors

$$\det(\alpha A + \beta B) \geq \det(A)^\alpha \det(B)^\beta. \quad (17.361) \quad \text{EqSPKooHFZymB}$$

Ici  $S^{++}$  est l'ensemble des matrices symétriques strictement définies positives, définition 8.229.

*Démonstration.* En plusieurs étapes.

43. La définition 15.91 du logarithme ne fonctionne que pour les réels strictement positifs. C'est le cas du déterminant d'une matrice réelle symétrique strictement définie positive.

- (1) **Pseudo-réduction** Le théorème de pseudo-réduction simultanée, corolaire 10.48, appliqué aux matrices  $A$  et  $B$  nous donne une matrice inversible  $Q$  telle que

$$\begin{cases} B = Q^t D Q \\ A = Q^t Q \end{cases} \quad \begin{matrix} (17.362a) \\ (17.362b) \end{matrix}$$

avec

$$D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \quad (17.363)$$

$\lambda_i > 0$ . Nous avons alors

$$\det(A)^\alpha \det(B)^\beta = \det(Q)^{2\alpha} \det(Q)^{2\beta} \det(D)^\beta = \det(Q)^2 \det(D)^\beta \quad (17.364)$$

(parce que  $\alpha + \beta = 1$ ) et

$$\det(\alpha A + \beta B) = \det(\alpha Q^t Q + \beta Q^t D Q) \quad (17.365a)$$

$$= \det(Q^t (\alpha \mathbb{1} + \beta D) Q) \quad (17.365b)$$

$$= \det(Q)^2 \det(\alpha \mathbb{1} + \beta D). \quad (17.365c)$$

- (2) **Ré-expression** L'inégalité (17.361) qu'il nous faut prouver se réduit donc à

$$\det(\alpha \mathbb{1} + \beta D) \geq \det(D)^\beta. \quad (17.366)$$

Vue la forme de  $D$  nous avons

$$\det(\alpha \mathbb{1} + \beta D) = \prod_{i=1}^n (\alpha + \beta \lambda_i) \quad (17.367)$$

et

$$\det(D)^\beta = \left( \prod_{i=1}^n \lambda_i \right)^\beta. \quad (17.368)$$

Il faut donc prouver que

$$\prod_{i=1}^n (\alpha + \beta \lambda_i) \geq \left( \prod_{i=1}^n \lambda_i \right)^\beta. \quad (17.369) \quad \text{EqGFLooQEicIS}$$

Cette dernière égalité de produit sera prouvée en passant au logarithme.

- (3) **Logarithme** Puisque le logarithme est concave par l'exemple 17.86, nous avons pour chaque  $i$  que

$$\ln(\alpha + \beta \lambda_i) \geq \alpha \ln(1) + \beta \ln(\lambda_i) = \beta \ln(\lambda_i). \quad (17.370)$$

En sommant cela sur  $i$  et en utilisant les propriétés de croissance et de multiplicativité du logarithme nous obtenons successivement

$$\sum_{i=1}^n \ln(\alpha + \beta \lambda_i) \geq \beta \sum_{i=1}^n \ln(\lambda_i) \quad (17.371a)$$

$$\ln \left( \prod_{i=1}^n (\alpha + \beta \lambda_i) \right) \geq \ln \left( \left( \prod_{i=1}^n \lambda_i \right)^\beta \right) \quad (17.371b)$$

$$\prod_{i=1}^n (\alpha + \beta \lambda_i) \geq \left( \prod_{i=1}^n \lambda_i \right)^\beta, \quad (17.371c)$$

ce qui est bien (17.369).

□

Rappel de notations :  $\mathbb{R}^+ = [0, \infty[$ . Voir la remarque 1.531.

**Lemme 17.93** ([1]).

Soit une fonction strictement convexe  $g: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ . Soit une fonction  $f: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$  vérifiant

- (1)  $f(0, 0) = 0$ ,
- (2)  $f(tx, ty) = tf(x, y)$  pour tout  $t$  (tant que ça ne déborde pas du domaine)
- (3)  $f(1, y) = g(y)$  pour tout  $y$
- (4)  $f(0, y) = f(y, 0)$ .

Alors  $f$  est convexe.

*Démonstration.* Nous devons prouver que pour toute paire de points  $A, B$  sur le graphe de  $f$ , le segment  $[A, B]$  est au-dessus du graphe de  $f$ . Ledit graphe étant d'ailleurs constitué de droites joignant  $(0, 0, 0)$  et les points du graphe de  $g$  (situé en  $x = 1$ ).

Nous notons  $\mathcal{C}$  le graphe de  $f$ .

- (1) **Une corde alignée à  $O$**  Soient deux point  $A$  et  $B$  alignés à l'origine  $O$ . Un point quelconque de  $[A, B]$  (et même de toute la droite) s'écrit

$$(tA_x, tA_y, tf(A_x, A_y)) = (tA_x, tA_y, f(tA_x, tA_y)), \quad (17.372)$$

et donc est sur le graphe de  $f$ .

- (2) **Autre corde** Nous prouvons que  $[A, B] \cap \mathcal{C} = \{A, B\}$ .

Si  $A$  et  $B$  sont dans le plan  $x = 1$  alors c'est d'accord parce que le graphe de  $f$  dans le plan est le même que celui de la fonction strictement convexe  $g$ .

Si  $A$  et  $B$  ne sont pas alignés à  $O$  et si ils ne sont pas dans le plan  $x = 1$  alors le plan  $AOB$  coupe le plan  $x = 1$  en une droite.

Nous supposons l'existence d'un point  $C \in ]A, B[ \cap \mathcal{C}$ .

Nous considérons la droite  $(OA)$  qui est contenue dans ce plan et dans  $\mathcal{C}$  (au moins la partie positive) et nous notons  $A'$  son intersection avec le plan  $x = 1$ . Même chose pour  $B$  et  $C$  qui donnent  $B'$  et  $C'$ .

Cela nous donne des points  $A', B'$  et  $C'$  qui sont alignés dans le graphe de  $f$  en  $x = 1$ . Or le graphe de  $f$  en  $x = 1$  est le graphe de la fonction  $g$  qui est strictement convexe et qui ne contient donc pas de points alignés.

Nous en concluons que si  $A, B \in \mathcal{C}$  alors  $]A, B[$  est soit complètement strictement au-dessus de  $\mathcal{C}$  soit complètement strictement en-dessous de  $\mathcal{C}$ .

Nous prouvons à présent que toutes les cordes sont au-dessus de  $\mathcal{C}$ . Pour cela, soient  $A, B \in ]0, \infty[ \times ]0, \infty[$ , deux points non alignés à  $O = (0, 0)$ . Nous considérons les points  $A', B'$  qui sont les intersections entre les droites  $(AO)$  et  $(BO)$  et la droite  $x = 1$  ainsi que le chemin  $\sigma$  qui parcourt le segment  $[A, A']$  et le chemin  $\gamma$  qui parcourt le segment  $[B, B']$  :

$$\begin{aligned} \sigma(0) &= A, & \gamma(0) &= B, \\ \sigma(1) &= A', & \gamma(1) &= B'. \end{aligned} \quad (17.373)$$

Pour tout  $u$ , la seule droite passant par  $O$  et par  $\sigma(u)$  passe également par  $A$ , et pas par  $B$ . En conséquence de quoi, pour tout  $u_1, u_2 \in [0, 1]$ , la droite  $(\sigma(u_1)\gamma(u_2))$  ne passe pas par  $(0, 0)$ .

Nous considérons à présent non seulement la corde joignant  $(A, f(A))$  à  $(B, f(B))$  et la corde joignant  $(A', f(A'))$  à  $(B', f(B'))$  mais également toutes les cordes intermédiaires (si vous aimez les grands mots, vous pouvez parler d'homotopie) :

$$c(u, t) = t(\sigma(u), f(\sigma(u))) + (1 - t)(\gamma(u), f(\gamma(u))) \quad (17.374)$$

Pour chaque  $u \in [0, 1]$ , cela représente une corde entre deux points non alignés à  $(0, 0, 0)$  et donc une corde qui est soit strictement au-dessus de  $\mathcal{C}$  soit strictement en-dessous (à par les points correspondant à  $t = 0$  et  $t = 1$  qui, eux, sont sur  $\mathcal{C}$ ).

Soit  $t_0 \in ]0, 1[$ . La courbe  $c(u, t_0)$  avec  $u \in [0, 1]$  ne touche jamais  $\mathcal{C}$ . Or le point  $c(1, t_0)$  est au-dessus de  $\mathcal{C}$ , donc le point  $c(0, t_0)$  est également au-dessus de  $\mathcal{C}$ .

Nous en concluons que toutes les cordes entre  $(A, f(A))$  et  $(B, f(B))$  est située au-dessus de  $\mathcal{C}$  et non en-dessous de  $\mathcal{C}$ .  $\square$

## 17.12.5 Convexité et hessienne

DEFooKCFPooLwKAsS

**Définition 17.94.**

Soit une partie convexe  $U$  de  $\mathbb{R}^n$  et une fonction  $f: U \rightarrow \mathbb{R}$ .

- (1) La fonction  $f$  est **convexe** si pour tout  $x, y \in U$  avec  $x \neq y$  et pour tout  $\theta \in ]0, 1[$  nous avons

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y). \quad (17.375)$$

- (2) Elle est **strictement convexe** si nous avons l'inégalité stricte.

PROPooYNNHooSHLvHp

**Proposition 17.95** ([576]).

Soit  $\Omega$  ouvert dans  $\mathbb{R}^n$  et  $U$  convexe dans  $\Omega$ , et une fonction différentiable  $f: U \rightarrow \mathbb{R}$ .

ITEMooRVIVooIayPS

- (1) La fonction  $f$  est convexe sur  $U$  si et seulement si pour tout  $x, y \in U$ ,

$$f(y) \geq f(x) + df_x(y - x). \quad (17.376)$$

ITEMooCWEWooFtNnKI

- (2) La fonction  $f$  est strictement convexe sur  $U$  si et seulement si pour tout  $x, y \in U$  avec  $x \neq y$ ,

$$f(y) > f(x) + df_x(y - x). \quad (17.377)$$

*Démonstration.* Nous avons quatre petites choses à démontrer.

- (1) **(1) sens direct** Soit une fonction convexe  $f$ . Nous avons :

$$f((1 - \theta)x + \theta y) \leq (1 - \theta)f(x) + \theta f(y), \quad (17.378)$$

donc

$$f(x + \theta(y - x)) - f(x) \leq \theta(f(y) - f(x)) \quad (17.379)$$

Puisque  $\theta > 0$  nous pouvons diviser par  $\theta$  sans changer le sens de l'inégalité :

$$\frac{f(x + \theta(y - x)) - f(x)}{\theta} \leq f(y) - f(x). \quad (17.380)$$

EQooAXXFooHWtiJh

Nous prenons la limite  $\theta \rightarrow 0^+$ . Cette limite est égale à la limite simple  $\theta \rightarrow 0$  et vaut (parce que  $f$  est différentiable) :

$$\frac{\partial f}{\partial(y - x)}(x) \leq f(y) - f(x), \quad (17.381)$$

et aussi

$$df_x(y - x) \leq f(y) - f(x) \quad (17.382)$$

par le lemme 12.280.

- (2) **(1) sens inverse** Pour tout  $a \neq b$  dans  $U$  nous avons

$$f(b) \geq f(a) + df_a(b - a). \quad (17.383)$$

EQooEALSpooJQszWr

Pour  $x \neq y$  dans  $U$  et pour  $\theta \in ]0, 1[$  nous écrivons (17.383) pour les couples  $(\theta x + (1 - \theta)y, y)$  et  $(\theta x + (1 - \theta)y, x)$ . Ça donne :

$$f(y) \geq f(\theta x + (1 - \theta)y) + df_{\theta x + (1 - \theta)y}(\theta(y - x)), \quad (17.384)$$

et

$$f(x) \geq f(\theta x + (1 - \theta)y) + df_{\theta x + (1 - \theta)y}((1 - \theta)(x - y)). \quad (17.385)$$

La différentielle est linéaire ; en multipliant la première par  $(1 - \theta)$  et la seconde par  $\theta$  et en la somme, les termes en  $df$  se simplifient et nous trouvons

$$\theta f(x) + (1 - \theta)f(y) \geq f(\theta x + (1 - \theta)y). \quad (17.386)$$

- (3) **(2) sens direct** Nous avons encore l'équation (17.380), avec une inégalité stricte. Par contre, ça ne va pas être suffisant parce que le passage à la limite ne conserve pas les inégalités strictes. Nous devons donc être plus malins.

Soient  $0 < \theta < \omega < 1$ . Nous avons  $(1 - \theta)x + \theta y \in [x, (1 - \omega)x + \omega y]$ , donc nous pouvons écrire  $(1 - \theta)x + \theta y$  sous la forme  $(1 - s)x + s((1 - \omega)x + \omega y)$ . Il se fait que c'est bon pour  $s = \theta/\omega$  (et aussi que nous avons  $\theta/\omega < 1$ ). Donc nous avons

$$f((1 - \theta)x + \theta y) = f\left((1 - \frac{\theta}{\omega})x + \frac{\theta}{\omega}((1 - \omega)x + \omega y)\right) \quad (17.387a)$$

$$< (1 - \frac{\theta}{\omega})f(x) + \frac{\theta}{\omega}f((1 - \omega)x + \omega y). \quad (17.387b)$$

Cela nous permet d'écrire

$$\frac{f((1 - \theta)x + \theta y) - f(x)}{\theta} < \frac{f((1 - \omega)x + \omega y) - f(x)}{\omega} < f(y) - f(x). \quad (17.388)$$

Le seconde inégalité est le pendant de (17.380). Maintenant en passant à la limite pour  $\theta$  nous conservons une inégalité stricte par rapport à  $f(y) - f(x)$  :

$$df_x(y - x) < f(y) - f(x). \quad (17.389)$$

□

Avant de lire la proposition suivante, il faut relire la proposition 12.375 et ce qui s'y rapporte. Lire aussi la remarque 17.85 qui indique qu'il n'y a pas de réciproque dans l'énoncé (2).

REMooYCRKooEQNikC

#### Remarque 17.96.

Notons que la condition (17.390) n'est pas équivalente à demander  $(d^2f)_x(h, h) \geq 0$  pour tout  $h$ . En effet nous ne demandons la positivité que dans les directions atteignables comme différence de deux éléments de  $U$ . La partie  $U$  n'est pas spécialement ouverte ; elle pourrait n'être qu'une droite dans  $\mathbb{R}^3$ . Dans ce cas, demander que  $f$  (qui est  $C^2$  sur l'ouvert  $\Omega$ ) soit convexe sur  $U$  ne demande que la positivité de  $(d^2f)_x$  appliqué à des vecteurs situés sur la droite  $U$ .

PROPooBMIRooFkQSAAb

#### Proposition 17.97 ([576]).

Soit une fonction  $f: \Omega \rightarrow \mathbb{R}$  deux fois différentiable sur l'ouvert  $\Omega$  de  $\mathbb{R}^n$  et un convexe  $U \subset \Omega$ .

ITEMooZQCAooIFjHOn

- (1) La fonction  $f$  est convexe sur  $U$  si et seulement si

$$(d^2f)_x(y - x, y - x) \geq 0 \quad (17.390)$$

EQooIBDCooJYdiRb

pour tout  $x, y \in U$ .

ITEMooHAGQooYZyhQk

- (2) Si pour tout  $x \neq y$  dans  $U$  nous avons

$$(d^2f)_x(y - x, y - x) > 0 \quad (17.391)$$

alors la fonction  $f$  est strictement convexe sur  $U$ .

*Démonstration.* Il y a trois parties à démontrer.

- (1) **(1) sens direct**

Soit une fonction convexe  $f$  sur  $U$ . Soient aussi  $x, y \in U$  et  $h = y - x$ . Nous utilisons ma version préférée de Taylor<sup>44</sup> : celle de la proposition 12.505 :

$$f(x + th) = f(x) + tdf_x(h) + \frac{t^2}{2}(d^2f_x)(h, h) + t^2\|h\|^2\alpha(th) \quad (17.392)$$

44. Si vous présentez ceci au jury d'un concours, vous devriez être capable de raconter ce que signifie  $d^2f$ , et pourquoi nous l'utilisons comme une 2-forme.

avec  $\lim_{s \rightarrow 0} \alpha(s) = 0$ . Le fait que  $f$  soit convexe donne

$$0 \leq f(x + th) - f(x) - tdf_x(h), \quad (17.393)$$

et donc

$$0 \leq \frac{t^2}{2} (d^2 f)_x(h, h) + f^2 \|h\|^2 \alpha(th). \quad (17.394)$$

En multipliant par 2 et en divisant par  $t^2$ ,

$$0 \leq (d^2 f)_x(h, h) + 2\|h\|^2 \alpha(th). \quad (17.395)$$

En prenant  $t \rightarrow 0$  nous avons bien  $(d^2 f)_x(y - x, y - x) \geq 0$ .

(2) **(1) sens inverse**

Soient  $x, y \in U$ . Nous écrivons Taylor en version de la proposition 12.506 :

$$f(y) = f(x) + df_x(y - x) + \frac{1}{2} (d^2 f)_z(y - x, y - x) \quad (17.396)$$

pour un certain  $z \in ]x, y[$ . En vertu de ce qui a été dit dans la remarque 17.96 nous ne pouvons pas évoquer l'hypothèse (17.390) pour conclure que  $(d^2 f)_z(y - x, y - x) \geq 0$ . Il y a deux manières de nous sortir du problème :

- Trouver  $s \in U$  tel que  $y - x = s - z$ .
- Trouver un  $z \in U$  tel que  $z - x$  soit un multiple de  $y - x$ .

La première approche ne fonctionne pas parce que  $s = y - x + z$  n'est pas garanti d'être dans  $U$  ; par exemple avec  $x = 1$ ,  $z = 2$ ,  $y = 3$  et  $U = [0, 3]$ . Dans ce cas  $s = 4 \notin U$ .

Pour suivre la seconde approche, nous posons  $z = \theta x + (1 - \theta)y$ , donc  $z - x = (1 - \theta)(y - x)$ . Dans ce cas la bilinéarité de  $(d^2 f)_z$  donne<sup>45</sup>

$$f(y) = f(x) + df_x(y - x) + \underbrace{\frac{1}{2} \frac{1}{(1 - \theta)^2} (d^2 f)_z(z - x, z - x)}_{\geq 0}. \quad (17.397)$$

Nous en déduisons que  $f$  est convexe par la proposition 17.95(1).

(3) **(2)**

Le raisonnement que nous venons de faire pour le sens inverse de (1) tient encore, et nous avons

$$f(y) = f(x) + df_x(y - x) + \underbrace{\frac{1}{2} \frac{1}{(1 - \theta)^2} (d^2 f)_z(z - x, z - x)}_{> 0} \quad (17.398)$$

d'où nous déduisons la stricte convexité de  $f$  par la proposition 17.95(2).

□

COR00MBQMooWBAlIH

**Corolaire 17.98.**

Soit un ouvert  $\Omega$  de  $\mathbb{R}^n$  et une fonction deux fois différentiable  $f$  sur  $\Omega$ .

ITEM00UAFTooXfCvii

(1) La fonction  $f$  est convexe si et seulement si pour tout  $x$ , la matrice hessienne  $d^2 f_x$  est semi-définie positive.

ITEM00DGISooPIRLod

(2) Si pour tout  $x$  de  $\Omega$ , la matrice hessienne  $d^2 f_x$  est strictement définie positive, alors  $f$  est strictement convexe.

45. Si vous avez bien suivi, la bilinéarité est contenue dans la proposition 12.375.

*Démonstration.* Nous pouvons voir ce résultat comme une conséquence directe de la proposition 17.97 en posant  $U = \Omega$ . Nous allons cependant en donner une démonstration directe.

Soit  $a \in \Omega$  et posons la fonction

$$\begin{aligned} g: \Omega &\rightarrow \mathbb{R} \\ x &\mapsto f(x) - f(a) - (df)_a(x - a). \end{aligned} \quad (17.399)$$

Nous allons calculer des différentielles de  $f$ , et une chose importante à comprendre est que la différentielle de la fonction  $x \mapsto df_a(x - a)$  ne fait pas intervenir la différentielle seconde de  $f$ ; c'est la différentielle de  $a \mapsto df_a(x)$  qui demanderait la différentielle seconde de  $f$ . Ici le point  $a$  étant donné,  $df_a$  est une application linéaire sans histoires. En particulier,  $df_a(x - a) = df_a(x) - df_a(a)$ .

La fonction  $g$  vérifie :

- (1)  $g(a) = 0$ ,
- (2)  $dg_x = df_x - df_a$ , parce que la différentielle de  $x \mapsto df_a(x)$  est  $x \mapsto df_a(x)$  en vertu du lemme 10.300.
- (3)  $dg_a = 0$ . Le point  $a$  est un point critique de  $g$ .
- (4)  $d^2g_x = d^2f_x$  parce que la différentielle de  $x \mapsto df_a$  est nulle.

Ceci étant dit, nous pouvons commencer avec la preuve.

(1) **(1) sens direct**

Nous supposons que  $f$  est convexe. Alors  $g(x) \geq 0$  pour tout  $x$  par la caractérisation 17.95(1). Cela signifie que  $x = 0$  est un minimum global de  $g$ . Par conséquent la proposition 17.70(3) nous dit que la Hessienne  $d^2f_a$  est semi-définie positive.

(2) **(1) sens inverse**

Nous sommes dans le cas de la proposition 17.70(1). Le point  $x = a$  est un minimum local de  $g$ , ce qui signifie que  $g(x) \geq 0$  pour tout  $x$  de  $\Omega$ . Encore une fois la caractérisation 17.95(1) nous permet de conclure.

(3) **(2)**

La fonction  $g$  vérifie les conditions de 17.70(2), donc  $x = 0$  est un minimum local strict de  $g$ . La caractérisation 17.95(2) nous fait conclure que  $f$  est strictement convexe.

□

### 17.99.

Nous rappelons que, avec  $p > 1$ , la fonction

$$\begin{aligned} f: [0, \infty[ &\rightarrow \mathbb{R} \\ x &\mapsto x^p \end{aligned} \quad (17.400)$$

est strictement croissante. La proposition 12.427 est formelle sur ce point.

### Lemme 17.100.

Soient un réel  $a$ , et  $p > 1$ . L'équation

$$|x|^p = ax \quad (17.401)$$

possède au plus deux solutions réelles.

*Démonstration.* En deux parties suivant le signe de  $a$ .

- (1) **Si  $a = 0$**  Alors l'équation est  $|x|^p = 0$ , et l'unique solution est  $x = 0$
- (2) **Si  $a \geq 0$**  Si  $x < 0$  alors  $ax < 0$  et  $|x|^p - ax > |x|^p > 0$ . Cela prouve que notre équation n'a pas de solutions  $x < 0$  lorsque  $a > 0$ .

Cherchons donc des solutions avec  $x \geq 0$ . D'abord  $|x| = x$  et ensuite, en posant  $p = 1 + \delta$  ( $\delta > 0$ ), nous avons la factorisation

$$x^p - ax = x(x^\delta - a) = 0. \quad (17.402)$$

Poser  $x = 0$  est clairement une solution. Si  $x \neq 0$ , nous avons le raisonnement suivant. Les réels formant un corps, c'est un anneau intègre qui vérifie alors la règle du produit nul<sup>46</sup>. Cela pour dire que si  $x \neq 0$ , alors

$$x^\delta = a. \quad (17.403)$$

La proposition 12.435 nous dit que cela a une unique solution dans les réels positifs.

(3) Si  $a < 0$  Ce cas se traite de façons similaire<sup>47</sup>.

□

PROPooLIGIooPrHYIb

### Proposition 17.101.

Soit  $1 < p < \infty$ . La fonction

$$\begin{aligned} f: \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto \|x\|^p \end{aligned} \quad (17.404)$$

est strictement convexe.

*Démonstration.* La preuve va se diviser en deux parties. D'abord nous allons utiliser la matrice Hessienne pour démontrer le résultat sur l'ouvert  $\mathbb{R}^n \setminus \{0\}$ , et ensuite nous allons un peu bricoler pour ajouter 0 au domaine de stricte convexité<sup>48</sup>.

(1) La Hessienne Nous notons

$$f_p(x) = \|x\|^p = \left( \sum_i x_i^2 \right)^{p/2}, \quad (17.405)$$

et nous dérivons :

$$\frac{\partial f_p}{\partial x_i}(x) = \frac{p}{2} 2x_i \left( \sum_i x_i^2 \right)^{(p-2)/2} = px_i f_{p-2}(x). \quad (17.406)$$

EQooNZZWooGpAlyj

Cela n'est déjà pas bien défini en  $x = 0$  lorsque  $p < 2$ , mais qu'importe ? Nous dérivons encore en utilisant entre autres la formule (17.406) elle-même avec  $p \rightarrow p - 2$  :

$$\frac{\partial^2 f_p}{\partial x_j \partial x_i}(x) = p\delta_{ij} \|x\|^{p-2} + p(p-2)x_i x_j \|x\|^{p-4}. \quad (17.407)$$

Nous avons donc la matrice Hessienne

$$H_{ij}(x) = p\delta_{ij} \|x\|^{p-2} + p(p-2)x_i x_j \|x\|^{p-4}. \quad (17.408)$$

Pour prouver que cette matrice est strictement définie positive, nous avons le choix entre la proposition 8.232(1) ou le lemme 8.233(1). Nous utilisons le second. Nous avons<sup>49</sup> :

$$y \cdot H(x)y = \sum_{kl} y_k H(x)_{kl} y_l \quad (17.409a)$$

$$= p \sum_l y_l^2 \|x\|^{p-2} + p(p-2) \|x\|^{p-4} \sum_{kl} x_k y_k x_l y_l \quad (17.409b)$$

$$= p \|y\|^2 \|x\|^{p-2} + p(p-2) \|x\|^{p-4} (x \cdot y)^2 \quad (17.409c)$$

$$= p \|x\|^{p-4} (\|y\|^2 \|x\|^2 + (p-2)(x \cdot y)^2) \quad (17.409d)$$

$$> p \|x\|^{p-4} (\|y\|^2 \|x\|^2 - (x \cdot y)^2) \quad (17.409e)$$

$$\geq 0. \quad (17.409f)$$

SUBEQooUSZQooCqgWPE

SUBEQooBXOKooZcarVv

Justifications :

— Pour 17.409e. Puisque  $p > 1$  nous avons  $p - 2 > -1$ . Là, l'inégalité est stricte et c'est important.

46. Voir le lemme 1.327 et la définition 1.326.

47. Vérifiez par vous-même et écrivez-moi si il y a un problème (personnellement, je n'ai pas vérifié.).

48. Si quelqu'un sait comment éviter ce bricolage en deux parties, je suis preneur.

49. Si vous ne savez pas où placer les indices, voyez la proposition 8.180.



— Pour 17.409f. C'est l'inégalité de Cauchy-Schwarz du théorème 10.1.

Voilà. La matrice Hessienne est strictement définie positive par le lemme 8.233(1) sur  $\mathbb{R}^n \setminus \{0\}$ . Le corolaire 17.98(1) nous indique que pour tout  $x, y \in \mathbb{R}^n \setminus \{0\}$  et pour tout  $\theta \in ]0, 1[$ ,

$$f_p(\theta x + (1 - \theta)y) < \theta f_p(x) + (1 - \theta)f_p(y) \quad (17.410)$$

pourvu que  $\theta x + (1 - \theta)y \neq 0$  pour tout  $\theta$ .

(2) **La suite**

Nous devons prouver que l'inéquation (17.410) tient également lorsque  $\theta x + (1 - \theta)y = 0$  pour une certaine valeur  $\theta = \theta_0 \in ]0, 1[$ .

(3) **Les cordes passant par zéro**

Pour ce faire, nous allons montrer que le segment de droite joignant  $(a, f_p(a))$  à  $(b, f_p(b))$  est toujours au-dessus de la courbe  $(x, f_p(x))$ . Nous commençons par  $b = 0$ . Vu que  $p > 1$  et que  $\theta \in ]0, 1[$  nous avons

$$\|\theta a\|^p = |\theta|^p \|a\|^p < \theta \|a\|^p = \theta f_p(a). \quad (17.411)$$

(4) **Les cordes passant au-dessus de zéro**

Dans le cas  $b \neq 0$  nous considérons

$$\begin{aligned} l_1: [a, b] &\rightarrow \mathbb{R} \\ x &\mapsto l_1(x) \end{aligned} \quad (17.412)$$

tel que  $(x, l_1(x))$  soit (le segment) la droite joignant  $(a, \|a\|^p)$  à  $(b, \|b\|^p)$ .

Nous considérons aussi, pour  $x \in [0, a]$  la fonction  $l_2$  telle que  $(x, l_2(x))$  soit la droite joignant  $(0, 0)$  à  $(a, \|a\|^p)$ . Nous avons déjà vu que pour  $x \in ]0, a[$  nous avons  $l_1(x) > \|x\|^p$ .

Nous avons  $l_1(a) = l_2(a) = \|a\|^p$ . Donc  $l_2(x) - l_1(x)$  ne change pas de signe sur  $[0, a]$ . Mais comme  $l_1(0) > 0 = l_2(0)$  nous avons

$$l_2(x) > l_1(x) \quad (17.413)$$

pour tout  $x \in [0, a]$ .

Au final,  $l_2(x) > l_1(x) > f_p(x)$ .

Pour la partie  $[b, 0]$  nous faisons de même en considérant  $l_3$  de telle sorte que  $(x, l_3(x))$  soit le segment joignant  $(0, 0)$  à  $(b, \|b\|^p)$ .

□

## 17.12.6 Quelques inégalités

### 17.12.6.1 Inégalité de Jensen

**Proposition 17.102** (Inégalité de Jensen).

Soit  $f: \mathbb{R} \rightarrow \mathbb{R}$  une fonction convexe et des réels  $x_1, \dots, x_n$ . Soient des nombres positifs  $\lambda_1, \dots, \lambda_n$  formant une combinaison convexe<sup>50</sup>. Alors

$$f\left(\sum_i \lambda_i x_i\right) \leq \sum_i \lambda_i f(x_i). \quad (17.414)$$

*Démonstration.* Nous procédons par récurrence sur  $n$ , en sachant que  $n = 2$  est la définition de la convexité de  $f$ . Vu que

$$\sum_{k=1}^n \lambda_k x_k = \lambda_n x_n + (1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k x_k}{1 - \lambda_n}, \quad (17.415)$$

50. Définition 11.31.

nous avons

$$f\left(\sum_{k=1}^n \lambda_k x_k\right) \leq \lambda_n f(x_n) + (1 - \lambda_n) f\left(\sum_{k=1}^{n-1} \frac{\lambda_k x_k}{1 - \lambda_n}\right). \quad (17.416)$$

La chose à remarquer est que les nombres  $\frac{\lambda_k}{1-\lambda_n}$  avec  $k$  allant de 1 à  $n-1$  forment eux-mêmes une combinaison convexe. L'hypothèse de récurrence peut donc s'appliquer au second terme du membre de droite :

$$f\left(\sum_{k=1}^n \lambda_k x_k\right) \leq \lambda_n f(x_n) + (1 - \lambda_n) \sum_{k=1}^{n-1} \frac{\lambda_k}{1 - \lambda_n} f(x_k) = \lambda_n f(x_n) + \sum_{k=1}^{n-1} \lambda_k f(x_k). \quad (17.417)$$

□

### 17.12.6.2 Inégalité arithmético-géométrique

La proposition suivante dit que la moyenne arithmétique de nombres strictement positifs est supérieure ou égale à la moyenne géométrique.

**Proposition 17.103** (Inégalité arithmético-géométrique[582]).

PropWDPooBtHIAr

Soient  $x_1, \dots, x_n$  des nombres strictement positifs. Nous posons

$$m_a = \frac{1}{n}(x_1 + \dots + x_n) \quad (17.418)$$

et

$$m_g = \sqrt[n]{x_1 \dots x_n} \quad (17.419)$$

Alors  $m_g \leq m_a$  et  $m_g = m_a$  si et seulement si  $x_i = x_j$  pour tout  $i, j$ .

*Démonstration.* Par hypothèse les nombres  $m_a$  et  $m_g$  sont tous deux strictement positifs, de telle sorte qu'il est équivalent de prouver  $\ln(m_g) \leq \ln(m_a)$  ou encore

$$\frac{1}{n}(\ln(x_1) + \dots + \ln(x_n)) \leq \ln\left(\frac{x_1 + \dots + x_n}{n}\right). \quad (17.420)$$

Cela n'est rien d'autre que l'inégalité de Jensen de la proposition 17.102 appliquée à la fonction  $\ln$  et aux coefficients  $\lambda_i = \frac{1}{n}$ . □

### 17.12.6.3 Inégalité de Kantorovitch

**Proposition 17.104** (Inégalité de Kantorovitch[583]).

PropMNUooFbYkug

Soit  $A$  une matrice symétrique strictement définie positive dont les plus grandes et plus petites valeurs propres sont  $\lambda_{\min}$  et  $\lambda_{\max}$ . Alors pour tout  $x \in \mathbb{R}^n$  nous avons

$$\langle Ax, x \rangle \langle A^{-1}x, x \rangle \leq \frac{1}{4} \left( \frac{\lambda_{\min}}{\lambda_{\max}} + \frac{\lambda_{\max}}{\lambda_{\min}} \right)^2 \|x\|^4. \quad (17.421)$$

*Démonstration.* Sans perte de généralité nous pouvons supposer que  $\|x\| = 1$ . Nous diagonalisons<sup>51</sup> la matrice  $A$  par la matrice orthogonale  $P \in O(n, \mathbb{R})$  :  $A = PDP^{-1}$  et  $A^{-1} = PD^{-1}P^{-1}$  où  $D$  est une matrice diagonale formée des valeurs propres de  $A$ .

Nous posons  $\alpha = \sqrt{\lambda_{\min}\lambda_{\max}}$  et nous regardons la matrice

$$\frac{1}{\alpha}A + \alpha A^{-1} \quad (17.422)$$

dont les valeurs propres sont

$$\frac{\lambda_i}{\alpha} + \frac{\alpha}{\lambda_i} \quad (17.423)$$

---

51. Théorème spectral 8.226.

parce que les vecteurs propres de  $A$  et de  $A^{-1}$  sont les mêmes (ce sont les valeurs de la diagonale de  $D$ ). Nous allons quelque peu étudier la fonction

$$\theta(x) = \frac{x}{\alpha} + \frac{\alpha}{x}. \quad (17.424)$$

Elle est convexe en tant que somme de deux fonctions convexes. Elle a son minimum en  $x = \alpha$  et ce minimum vaut  $\theta(\alpha) = 2$ . De plus

$$\theta(\lambda_{\max}) = \theta(\lambda_{\min}) = \sqrt{\frac{\lambda_{\min}}{\lambda_{\max}}} + \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}. \quad (17.425)$$

Une fonction convexe passant deux fois par la même valeur doit forcément être plus petite que cette valeur entre les deux<sup>52</sup> : pour tout  $x \in [\lambda_{\min}, \lambda_{\max}]$ ,

$$\theta(x) \leq \sqrt{\frac{\lambda_{\min}}{\lambda_{\max}}} + \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}. \quad (17.426)$$

Nous sommes maintenant en mesure de nous lancer dans l'inégalité de Kantorovitch.

$$\sqrt{\langle Ax, x \rangle \langle A^{-1}x, x \rangle} \leq \frac{1}{2} \left( \frac{\langle Ax, x \rangle}{\alpha} + \alpha \langle A^{-1}x, x \rangle \right) \quad (17.427a)$$

$$= \frac{1}{2} \left\langle \left( \frac{A}{\alpha} + \alpha A^{-1} \right) x, x \right\rangle \quad (17.427b)$$

$$\leq \frac{1}{2} \left\| \left( \frac{A}{\alpha} + \alpha A^{-1} \right) x \right\| \|x\| \quad (17.427c)$$

$$\leq \frac{1}{2} \left\| \frac{A}{\alpha} + \alpha A^{-1} \right\| \|x\|^2 \quad (17.427d)$$

Justifications :

- 17.427a par l'inégalité arithmético-géométrique, proposition 17.103. Nous avons aussi inséré  $\alpha \frac{1}{\alpha}$  dans le produit sous la racine.
- 17.427c par l'inégalité de Cauchy-Schwarz, théorème 10.1.
- 17.427d par la définition de la norme opérateur de la proposition 10.63

La norme opérateur est la plus grande des valeurs propres. Mais les valeurs propres de  $A/\alpha + \alpha A^{-1}$  sont de la forme  $\theta(\lambda_i)$ , et tous les  $\lambda_i$  sont entre  $\lambda_{\min}$  et  $\lambda_{\max}$ . Donc la plus grande valeur propre de  $A/\alpha + \alpha A^{-1}$  est  $\theta(x)$  pour un certain  $x \in [\lambda_{\min}, \lambda_{\max}]$ . Par conséquent

$$\sqrt{\langle Ax, x \rangle \langle A^{-1}x, x \rangle} \leq \frac{1}{2} \left\| \frac{A}{\alpha} + \alpha A^{-1} \right\| \|x\|^2 \leq \sqrt{\frac{\lambda_{\min}}{\lambda_{\max}}} + \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}}. \quad (17.428)$$

□

### 17.12.7 Norme $l^p$

Si vous êtes seulement intéressés par le cas  $p = 1, 2$  ou le cas  $p = \infty$ , allez voir la proposition 10.52.

#### Proposition 17.105.

Pour tout réel  $p \geq 1$ , la formule

$$\|x\|_{L^p} = \left( \sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad (17.429)$$

définit une norme<sup>53</sup> sur  $\mathbb{R}^n$ .

52. Je ne suis pas certain que cette phrase soit claire, non ?

53. Définition 7.212.

*Démonstration.* Le seul point un peu délicat parmi les propriétés à vérifier de la définition 7.212 est l'inégalité triangulaire. Grâce au lemme 7.214, nous allons seulement montrer que la partie  $B = \{x \in \mathbb{R}^n \text{ tel que } \|x\|_p \leq 1\}$  est convexe.

Soient  $x, y \in B$  ainsi que  $\lambda \in [0, 1]$ . La fonction

$$\alpha: [0, \infty[ \rightarrow [0, \infty[ \\ t \mapsto t^p \quad (17.430)$$

est convexe par le lemme 17.87. Nous avons donc le calcul suivant :

$$\begin{aligned} \|\lambda x + (1 - \lambda)y\|_p^p &= \sum_{i=1}^n |\lambda x_i + (1 - \lambda)y_i|^p & (17.431a) \\ &\leq \sum_i (\lambda|x_i| + (1 - \lambda)|y_i|)^p & \text{SUBEQooYETKooSSCgMa} \quad (17.431b) \\ &\leq \sum_i \lambda^p |x_i|^p + (1 - \lambda)^p |y_i|^p & \text{SUBEQooHNZHooKmWdJS} \quad (17.431c) \\ &\leq \sum_i \lambda |x_i|^p + \sum_i (1 - \lambda) |y_i|^p & \text{SUBEQooMUPCooHicEC} \quad (17.431d) \\ &= \lambda \|x\|_p^p + (1 - \lambda) \|y\|_p^p & (17.431e) \\ &\leq \lambda + (1 - \lambda) & \text{SUBEQooEFCHooJmSdVv} \quad (17.431f) \\ &= 1. & (17.431g) \end{aligned}$$

Justifications.

- Pour (17.431b). L'inégalité  $|a + b| \leq |a| + |b|$ .
- Pour (17.431c). La fonction  $\alpha$  est convexe.
- Pour (17.431d). Nous avons  $\lambda \in [0, 1]$  et  $p \geq 1$  et donc  $\lambda^p \leq \lambda$  par 12.428.
- Pour (17.431f). Parce que  $x \in B$  implique  $\|x\|_p \leq 1$  qui implique  $\|x\|_p^p \leq 1$ .

□

### 17.12.8 Hölder

Notre étude de la fonction puissance permet de démontrer quelques inégalités de Hölder. Voir le thème 25.

THOooPPDPooJxTYIy

**Théorème 17.106** ([584]).

Si  $1 \leq p < \infty$ , alors pour tout  $x \in \mathbb{R}^n$  nous avons<sup>54</sup>

$$\|x\|_\infty \leq \|x\|_p \leq n^{1/p} \|x\|_\infty. \quad (17.432)$$

*Démonstration.* Vu que la fonction  $t \mapsto |t|^p$  est croissante pour les  $t$  positifs<sup>55</sup>, pour chaque  $i$  nous avons

$$|x_i| \leq \left( \sum_k |x_k|^p \right)^{1/p} = \|x\|_p. \quad (17.433)$$

Cela montre que

$$\|x\|_\infty = \max\{|x_i|\} \leq \|x\|_p. \quad (17.434)$$

D'autre part, pour chaque  $i$  nous avons  $|x_i| \leq \|x\|_\infty$ , donc

$$\|x\|_p \leq (n \|x\|_\infty^p)^{1/p} = n^{1/p} \|x\|_\infty. \quad (17.435)$$

□

54. Définition de la norme  $\|\cdot\|_p$  sur  $\mathbb{R}^n$ , proposition 17.105.

55. Parce que  $p \geq 1$  et la proposition 12.427.

Le corolaire suivant donne une façon de majorer une norme  $\ell^p$  par une norme  $\ell^q$  moyennant un facteur. Notons cependant que l'inégalité de Hölder de la proposition 17.111 est plus précise. Ce corolaire est suffisant pour prouver l'équivalence des normes  $\ell^p$ .

CORooEZGHooACHOiB

**Corolaire 17.107.**

Soient  $p \geq 1$  et  $q \leq \infty$ . Pour tout  $x \in \mathbb{R}^n$  nous avons

$$\|x\|_p \leq n^{1/p} \|x\|_q. \quad (17.436)$$

*Démonstration.* Il suffit d'utiliser les deux inégalités du théorème 17.106. D'abord la seconde avec  $p$ , et ensuite la première avec  $q$ .  $\square$

Si vous cherchez l'inégalité de Hölder dans  $L^p$ , c'est la proposition 27.39. Les normes  $\ell^p$  sont définies dans 17.105.

LEMooLGGDooGLGFHj

**Lemme 17.108 ([585]).**

Soient  $x, y > 0$  ainsi que  $\alpha, \beta > 0$  tels que  $\alpha + \beta = 1$ . Alors

$$xy \leq \alpha x^{1/\alpha} + \beta y^{1/\beta}. \quad (17.437)$$

Nous avons

$$xy = \alpha x^{1/\alpha} + \beta y^{1/\beta} \quad (17.438)$$

si et seulement si  $x^{1/\alpha} = y^{1/\beta}$ .

*Démonstration.* Nous utilisons le logarithme<sup>56</sup> et ses propriétés (surtout la proposition 15.94). D'abord

$$xy = e^{\ln(xy)} = e^{\ln(x) + \ln(y)} = e^{\alpha \frac{\ln(x)}{\alpha} + \beta \frac{\ln(y)}{\beta}} \quad (17.439)$$

Comme l'exponentielle est strictement convexe (exemple 17.86(3)) et vu que  $\alpha + \beta = 1$ , nous avons

$$xy = e^{\alpha \frac{\ln(x)}{\alpha} + \beta \frac{\ln(y)}{\beta}} \leq \alpha e^{\ln(x)/\alpha} + \beta e^{\ln(y)/\beta} = \alpha x^{1/\alpha} + \beta y^{1/\beta}. \quad (17.440)$$

EQooNLQJooAYiEAO

Cela prouve déjà l'inégalité demandée.

Nous passons au cas d'égalité. Puisque  $\alpha$  et  $\beta$  ne sont pas nuls, l'inégalité (17.440) est une égalité si et seulement si<sup>57</sup>

$$\frac{\ln(x)}{\alpha} = \frac{\ln(y)}{\beta}. \quad (17.441)$$

Cela signifie  $\ln(x^{1/\alpha}) = \ln(y^{1/\beta})$ , qui implique  $x^{1/\alpha} = y^{1/\beta}$  parce que le logarithme est une bijection.  $\square$

CORooTCBZooAcZxaC

**Corolaire 17.109.**

Si  $p, q > 0$  vérifient  $\frac{1}{p} + \frac{1}{q} = 1$  et si  $p > 1$  alors nous avons

$$xy \leq \frac{1}{p} x^p + \frac{1}{q} y^q \quad (17.442)$$

EQooWKTSooQwRsLz

pour tout  $x, y \geq 0$ , avec une égalité si et seulement si  $x^p = y^q$ .

*Démonstration.* Il suffit de poser  $\alpha = 1/p$  et  $\beta = 1/q$  et appliquer le lemme 17.108.  $\square$

THOooYHMJooBlXfpl

**Théorème 17.110 (Inégalité de Hölder[585]).**

Soient  $p, q > 0$  tels que  $\frac{1}{p} + \frac{1}{q} = 1$ . Pour tout  $x, y \in \mathbb{R}^n$  nous avons<sup>58</sup>

$$|x \cdot y| = \sum_{i=1}^n |x_i y_i| \leq \|x\|_p \|y\|_q. \quad (17.443)$$

56. Définition 15.91

57. Par le lemme 17.79.

58. Définition de la norme  $\|\cdot\|_p$ , proposition 17.105.

Il y a égalité si et seulement si  $x_i y_i$  est de signe constant<sup>59</sup> et les vecteurs  $\sum_i |x_i|^p e_i$  et  $\sum_i |y_i|^q e_i$  sont proportionnels.

*Démonstration.* En plusieurs parties.

- (1) **Le cas des vecteurs nuls** Si  $x$  ou  $y$  est nul, les inégalités sont évidentes. Donc nous supposons que non.
- (2) **Première inégalité** En ce qui concerne la première inégalité,

$$|x \cdot y| = \left| \sum_i x_i y_i \right| \leq \sum_i |x_i| |y_i| \leq \sum_i \left( \frac{1}{p} |x_i|^p + \frac{1}{q} |y_i|^q \right) \quad (17.444)$$

où nous avons utilisé le corolaire 17.109 dans chaque terme de la somme en tenant compte du fait que  $|x_i|$  et  $|y_i|$  sont positifs.

- (3) **Seconde inégalité** Pour la seconde inégalité, nous commençons avec  $\|x\|_p = \|y\|_q = 1$ . Utilisant encore le corolaire 17.109 pour chaque terme, nous avons

$$\sum_i |x_i y_i| \leq \underbrace{\frac{1}{p} \sum_i |x_i|^p}_{=1} + \underbrace{\frac{1}{q} \sum_i |y_i|^q}_{=1} = \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p \|y\|_q. \quad (17.445)$$

Si maintenant  $x$  et  $y$  sont arbitraires non nuls dans  $\mathbb{R}^n$ , nous posons  $x' = x/\|x\|_p$  et  $y' = y/\|y\|_q$ ; nous savons déjà que

$$\sum_i |x'_i y'_i| \leq 1. \quad (17.446)$$

En remplaçant  $x'_i$  par  $x_i/\|x\|_p$  et  $y'_i$  par  $y_i/\|y\|_q$ , l'inégalité (17.446) devient

$$\sum_i \frac{|x_i| |y_i|}{\|x\|_p \|y\|_q} \leq 1, \quad (17.447)$$

ce qui signifie

$$\sum_i |x_i y_i| \leq \|x\|_p \|y\|_q. \quad (17.448)$$

- (4) **Cas d'égalité, dans un sens** Nous notons

$$\mathcal{D} = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \text{ tel que } \sum_i |x_i y_i| = \|x\|_p \|y\|_q\}. \quad (17.449)$$

- (4a) **Multiplications** D'abord si  $(x, y) \in \mathcal{D}$ , alors  $(\mu x, \lambda y) \in \mathcal{D}$  pour tout  $\mu, \lambda \in \mathbb{R}$ . En effet,

$$\sum_i |(\mu x)_i (\lambda y)_i| = |\mu| |\lambda| \sum_i |x_i y_i| = |\mu| |\lambda| \|x\|_p \|y\|_q = \|\mu x\|_p \|\lambda y\|_q. \quad (17.450)$$

- (4b) **Avec normes égales à 1** Soit  $(x, y) \in \mathcal{D}$  tels que  $\|x\|_p = \|y\|_q = 1$ . Nous avons en particulier,

$$1 = \sum_i |x_i y_i| \leq \frac{1}{p} \sum_i |x_i|^p + \frac{1}{q} \sum_i |y_i|^q \quad (17.451)$$

grâce à l'inégalité (17.442) appliquée à chaque terme. Vu que  $\|x\|_p = 1$ , nous avons  $\sum_i |x_i|^p = 1$ , de telle sorte que le membre de droite de (17.451) se réduit à  $\frac{1}{p} + \frac{1}{q} = 1$ .

Nous pouvons donc écrire

$$1 = \sum_i |x_i y_i| \leq \frac{1}{p} \sum_i |x_i|^p + \frac{1}{q} \sum_i |y_i|^q = 1. \quad (17.452)$$

59. Je n'utilise pas cette hypothèse de signe constant. Il doit y avoir une subtilité qui m'a échappée. Soyez prudente en lisant et écrivez-moi si vous trouvez une erreur.

L'inégalité est donc une égalité :

$$\sum_i |x_i y_i| = \sum_i \left( \frac{1}{p} |x_i|^p + \frac{1}{q} |y_i|^q \right). \quad (17.453)$$

Mais chaque terme à gauche est en inégalité avec le terme correspondant à droite :

$$|x_i y_i| \leq \frac{1}{p} |x_i|^p + \frac{1}{q} |y_i|^q. \quad (17.454)$$

Pour que le tout soit une égalité, il faut que chaque inégalité (17.454) soit une égalité. Pour chaque  $i$ , nous avons

$$|x_i y_i| = \frac{1}{p} |x_i|^p + \frac{1}{q} |y_i|^q. \quad (17.455)$$

La condition d'égalité du corolaire 17.109 nous dit alors que  $|x_i|^p = |y_i|^q$ .

- (4c) **Avec normes arbitraires** Soit donc  $(x, y) \in \mathcal{D}$ . Nous savons qu'en posant  $x' = x/\|x\|_p$  et  $y' = y/\|y\|_q$  nous avons  $(x', y') \in \mathcal{D}$  et donc

$$\left( \frac{|x_i|}{\|x\|_p} \right)^p = \left( \frac{|y_i|}{\|y\|_q} \right)^q. \quad (17.456)$$

Cela donne tout de suite

$$|x_i|^p = \frac{\|x\|_p^p}{\|y\|_q^q} |y_i|^q, \quad (17.457)$$

ce qui est bien ce que nous voulions : le vecteur  $\sum_i |x_i|^p e_i$  est proportionnel au vecteur  $\sum_i |y_i|^q e_i$ .

- (5) **Cas d'égalité dans l'autre sens** [1] Nous supposons que les vecteurs  $\sum_i |x_i|^p e_i$  et  $\sum_i |y_i|^q e_i$  sont proportionnels. Nous nommons  $c^q$  le facteur de proportionnalité, c'est-à-dire que nous posons

$$|x_i|^p = c^q |y_i|^q. \quad (17.458)$$

Dans ce cas, pour chaque  $i$ , les nombres  $c|y_i|$  et  $|x_i|$  sont dans le cas d'égalité du corolaire 17.109. Nous avons alors

$$\sum_i |x_i y_i| = \frac{1}{c} \sum_i c |x_i| |y_i| \quad (17.459a)$$

$$= \frac{1}{c} \sum_i \left( \frac{1}{p} |x_i|^p + \frac{1}{q} (c|y_i|)^q \right) \quad (17.459b)$$

$$= \frac{1}{c} \sum_i \left( \frac{1}{p} |x_i|^p + \frac{1}{q} |x_i|^p \right) \quad (17.459c)$$

$$= \frac{1}{c} \sum_i |x_i|^p. \quad (17.459d)$$

Et c'est maintenant que nous subdivisons.

- (5a) **Si  $\|x\|_p = \|y\|_q = 1$**  Dans ce cas, l'égalité (17.459) se réduit à

$$\sum_i |x_i y_i| = \frac{1}{c}. \quad (17.460)$$

Mais l'hypothèse sur les normes donne

$$1 = \sum_i |x_i|^p = \sum_i c^q |y_i|^q = c^q \sum_i |y_i|^q = c^q. \quad (17.461)$$

Donc  $c = 1$  et nous avons bien

$$\sum_i |x_i y_i| = \frac{1}{c} = 1 = \|x\|_p \|y\|_q. \quad (17.462)$$

- (5b) **Pour des normes arbitraires** Soit  $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$  tels que  $|x_i|^p = c^q |y_i|^q$ . Nous posons comme d'habitude  $x' = x/\|x\|_p$  et  $y' = y/\|y\|_q$ . En utilisant le cas « de norme 1 » nous avons

$$1 = \sum_i |x'_i y'_i| = \frac{1}{\|x\|_p \|y\|_q} \sum_i |x_i y_i|. \quad (17.463)$$

Donc  $\sum_i |x_i y_i| = \|x\|_p \|y\|_q$  comme nous le voulions.

□

La majoration de la proposition suivante sera utile pour les inégalités de Clarkson du lemme 27.146. Pour d'autres inégalités (plus simples) autour des normes  $\|\cdot\|_p$ , voir le thème 25.

PROPooQZTNooGACMIQ

**Proposition 17.111** ([585]).

Si  $x \in \mathbb{R}^n$  et si  $0 < q < p$ , alors

$$\|x\|_q \leq n^{\frac{1}{q} - \frac{1}{p}} \|x\|_p. \quad (17.464)$$

En particulier, si  $0 < q < p$ , alors

$$\|x\|_q \leq \|x\|_p. \quad (17.465)$$

*Démonstration.* Nous posons  $P = p/q$  et  $Q = P/(P-1)$ . Les nombres  $P$  et  $Q$  sont des exposants conjugués, parce que

$$\frac{1}{P} + \frac{1}{Q} = \frac{q}{p} + \frac{p-q}{p} = 1. \quad (17.466)$$

Nous posons  $y = (1, \dots, 1) \in \mathbb{R}^n$  ainsi que

$$v = \sum_i |x_i|^q e_i, \quad (17.467)$$

et nous écrivons l'inégalité de Hölder de la proposition 17.110 sur les vecteurs  $v$  et  $y$  :

$$\sum_i |v_i y_i| \leq \|v\|_P \|y\|_Q. \quad (17.468)$$

En déballant,

$$\sum_i |x_i|^q \leq \left( \sum_i (|x_i|^q)^P \right)^{1/P} \underbrace{\left( \sum_i 1^Q \right)^{1/Q}}_{=n} \quad (17.469a)$$

$$= \left( \sum_i |x_i|^p \right)^{q/p} n^{1/Q} \quad (17.469b)$$

$$= n^{1-q/p} \|x\|_p^q. \quad (17.469c)$$

Cela donne

$$\|x\|_q^q \leq n^{1-q/p} \|x\|_p^q. \quad (17.470)$$

En prenant la puissance  $1/q$  des deux côtés,

$$\|x\|_q \leq n^{\frac{1}{q} - \frac{1}{p}} \|x\|_p \leq \|x\|_p \quad (17.471)$$

parce que  $n \geq 1$  et  $p/q < 1$ .

□



## 17.13 Algorithme du gradient à pas optimal

Une idée pour trouver un minimum à une fonction est de prendre un point  $p$  au hasard, calculer le gradient  $\nabla f(p)$  et suivre la direction  $-\nabla f(p)$  tant que ça descend. Une fois qu'on est « dans le creux », recalculer le gradient et continuer ainsi.

Nous allons détailler cet algorithme dans un cas très particulier d'une matrice  $A$  symétrique et strictement définie positive.

- Dans la proposition 17.113 nous montrons que résoudre le système linéaire  $Ax = -b$  est équivalent à minimiser une certaine fonction.
- La proposition 17.114 donnera une méthode itérative pour trouver ce minimum.

DefQXPooYSygGP

### Définition 17.112.

Si  $X$  est un espace vectoriel normé et  $f: X \rightarrow \mathbb{R} \cup \{\pm\infty\}$  nous disons que  $f$  est **coercive** sur le domaine non borné  $P$  de  $X$  si pour tout  $M \in \mathbb{R}$ , l'ensemble

$$\{x \in P \text{ tel que } f(x) \leq M\} \quad (17.472)$$

est borné.

En langage imagé la coercivité de  $f$  s'exprime par la limite

$$\lim_{\substack{\|x\| \rightarrow \infty \\ x \in P}} f(x) = +\infty. \quad (17.473)$$

Nous rappelons que  $S^{++}(n, \mathbb{R})$  est l'ensemble des matrices symétriques strictement définies positives définies en 8.230.

PROPooYRLDooTwzfWU

### Proposition 17.113.

Soit  $A \in S^{++}(n, \mathbb{R})$  et  $b \in \mathbb{R}^n$ . Nous considérons l'application

$$\begin{aligned} f: \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle. \end{aligned} \quad (17.474)$$

Alors :

- (1) Il existe un unique  $\bar{x} \in \mathbb{R}^n$  tel que  $A\bar{x} = -b$ .
- (2) Il existe un unique  $x^* \in \mathbb{R}^n$  minimisant  $f$ .
- (3) Ils sont égaux :  $\bar{x} = x^*$ .

*Démonstration.* Une matrice symétrique strictement définie positive est inversible, entre autres parce qu'elle se diagonalise par des matrices orthogonales (qui sont inversibles) et que la matrice diagonalisée est de déterminant non nul : tous les éléments diagonaux sont strictement positifs. Voir le théorème spectral symétrique 8.226.

D'où l'unicité du  $\bar{x}$  résolvant le système  $Ax = -b$  pour n'importe quel  $b$ .

#### (1) $f$ est strictement convexe

La fonction  $f$  s'écrit

$$f(x) = \frac{1}{2} \sum_{kl} A_{kl} x_l x_k + \sum_k b_k x_k. \quad (17.475)$$

Elle est de classe  $C^2$  sans problèmes, et il est vite vu que  $\frac{\partial^2 f}{\partial x_i \partial x_j} = A_{ij}$ , c'est-à-dire que  $A$  est la matrice hessienne de  $f$ . Cette matrice étant strictement définie positive par hypothèse, la fonction  $f$  est strictement convexe par le corolaire 17.98(2).

(2)  $f$  est coercive Montrons à présent que  $f$  est coercive. Nous avons :

$$|f(x)| = \left| \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle \right| \quad (17.476a)$$

$$\geq \frac{1}{2} |\langle Ax, x \rangle| - |\langle b, x \rangle| \quad (17.476b)$$

$$\geq \frac{1}{2} \lambda_{\min} \|x\|^2 - \|b\| \|x\| \quad (17.476c)$$

Pour la dernière ligne nous avons nommé  $\lambda_{\min}$  la plus grande valeur propre de  $A$  et utilisé Cauchy-Schwarz pour le second terme. Nous avons donc bien  $|f(x)| \rightarrow \infty$  lorsque  $\|x\| \rightarrow \infty$  et la fonction  $f$  est coercive.

Soit  $M$  une valeur atteinte par  $f$ . L'ensemble

$$\{x \in \mathbb{R}^n \text{ tel que } f(x) \leq M\} \quad (17.477)$$

est fermé (parce que  $f$  est continue) et borné parce que  $f$  est coercive. Cela est donc compact<sup>60</sup> et  $f$  atteint un minimum qui sera forcément dedans. Cela est pour l'existence d'un minimum.

Pour l'unicité du minimum nous invoquons la convexité : si  $\bar{x}_1$  et  $\bar{x}_2$  sont deux points réalisant le minimum de  $f$ , alors

$$f\left(\frac{\bar{x}_1 + \bar{x}_2}{2}\right) < \frac{1}{2}f(\bar{x}_1) + \frac{1}{2}f(\bar{x}_2) = f(\bar{x}_1), \quad (17.478)$$

ce qui contredit la minimalité de  $f(\bar{x}_1)$ .

Nous devons maintenant prouver que  $\bar{x}$  vérifie l'équation  $A\bar{x} = -b$ . Vu que  $\bar{x}$  est minimum local de  $f$  qui est une fonction de classe  $C^2$ , le théorème des minimums locaux 17.69 nous indique que  $\bar{x}$  est solution de  $\nabla f(x) = 0$ . Calculons un peu cela avec la formule

$$df_x(u) = \frac{d}{dt} \left[ f(x+tu) \right]_{t=0} = \frac{1}{2} (\langle Ax, u \rangle + \langle Au, x \rangle) + \langle b, u \rangle = \langle Ax, u \rangle + \langle b, u \rangle = \langle Ax+b, u \rangle. \quad (17.479)$$

Donc demander  $df_x(u) = 0$  pour tout  $u$  demande  $Ax + b = 0$ . □

PropSOOooGoMOxG

**Proposition 17.114** (Gradient à pas optimal).

Soit  $A \in S^{++}(n, \mathbb{R})$  ( $A$  est une matrice symétrique strictement définie positive) et  $b \in \mathbb{R}^n$ . Nous considérons l'application

$$\begin{aligned} f: \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{2} \langle Ax, x \rangle + \langle b, x \rangle. \end{aligned} \quad (17.480)$$

Soit  $x_0 \in \mathbb{R}^n$ . Nous définissons la suite  $(x_k)$  par

$$x_{k+1} = x_k + t_k d_k \quad (17.481)$$

où

- $d_k = -(\nabla f)(x_k)$
- $t_k$  est la valeur minimisant la fonction  $t \mapsto f(x_k + td_k)$  sur  $\mathbb{R}$ .

Alors pour tout  $k \geq 0$  nous avons

$$\|x_k - \bar{x}\| \leq K \left( \frac{c_2(A) - 1}{c_2(A) + 1} \right)^k \quad (17.482)$$

où  $c_2(A) = \frac{\lambda_{\max}}{\lambda_{\min}}$  est le rapport entre la plus grande et la plus petite valeur propre<sup>61</sup> de la matrice  $A$  et  $\bar{x}$  est l'unique élément de  $\mathbb{R}^n$  à minimiser  $f$ .

60. Théorème 9.24

61. Cela est certainement très lié au conditionnement de la matrice  $A$ , voir la proposition 34.110.

*Démonstration.* Décomposition en plusieurs points.

- (1) **Existence de  $\bar{x}$**  Le fait que  $\bar{x}$  existe et soit unique est la proposition 17.113.
- (2) **Si  $(\nabla f)(x_k) = 0$**  D'abord si  $\nabla f(x_k) = 0$ , c'est que  $x_{k+1} = x_k$  et l'algorithme est terminé : la suite est stationnaire. Pour dire que c'est gagné, nous devons prouver que  $x_k = \bar{x}$ . Pour cela nous écrivons (à partir de maintenant «  $x_k$  » est la  $k^{\text{e}}$  composante de  $x$  qui est une variable, et non le  $x_k$  de la suite)

$$f(x) = \frac{1}{2} \sum_{kl} A_{kl} x_l x_k + \sum_k b_k x_k \quad (17.483)$$

et nous calculons  $\frac{\partial f}{\partial x_i}(a)$  en tenant compte du fait que  $\frac{\partial x_k}{\partial x_i} = \delta_{ki}$ . Le résultat est que  $(\partial_i f)(a) = (Ax + b)_i$  et donc que

$$(\nabla f)(a) = Aa + b. \quad (17.484)$$

Vu que  $A$  est inversible (symétrique définie positive), il existe un unique  $a \in \mathbb{R}^n$  qui vérifie cette relation. Par la proposition 17.113, cet élément est le minimum  $\bar{x}$ .

Cela pour dire que si  $a \in \mathbb{R}^n$  vérifie  $(\nabla f)(a) = 0$  alors  $a = \bar{x}$ . Nous supposons donc à partir de maintenant que  $\nabla f(x_k) \neq 0$  pour tout  $k$ .

- (3)  **$t_k$  est bien défini**

Pour  $t \in \mathbb{R}$  nous avons

$$f(x_k + td_k) = f(x_k) + \frac{1}{2}t^2 \langle Ad_k, d_k \rangle + t \underbrace{\langle Ax_k + b, d_k \rangle}_{=-d_k} = \frac{1}{2}t^2 \langle Ad_k, d_k \rangle - t \|d_k\|^2 + f(x_k). \quad (17.485)$$

qui est un polynôme du second degré en  $t$ . Le coefficient de  $t^2$  est  $\frac{1}{2} \langle Ad_k, d_k \rangle > 0$  parce que  $d_k \neq 0$  et  $A$  est strictement définie positive. Par conséquent la fonction  $t \mapsto f(x_k + td_k)$  admet bien un unique minimum. Nous pouvons même calculer  $t_k$  parce que l'on connaît par cœur le sommet d'une parabole :

$$t_k = -\frac{\langle Ax_k + b, d_k \rangle}{\langle Ad_k, d_k \rangle} = \frac{\|d_k\|^2}{\langle Ad_k, d_k \rangle} \quad (17.486)$$

parce que  $d_k = -\nabla f(x_k) = -(Ax_k + b)$ .

- (4) **La valeur de  $d_{k+1}$**

Par définition,  $d_{k+1} = -\nabla f(x_{k+1}) = -(Ax_{k+1} + b)$ . Mais  $x_{k+1} = x_k + t_k d_k$ , donc

$$d_{k+1} = -Ax_k - t_k Ad_k - b = d_k - t_k Ad_k \quad (17.487)$$

parce que  $-Ax_k - b = d_k$ .

Par ailleurs,  $\langle d_{k+1}, d_k \rangle = 0$  parce que

$$\langle d_{k+1}, d_k \rangle = \langle d_k, d_k \rangle - t_k \langle d_k, Ad_k \rangle = \|d_k\|^2 - \frac{\|d_k\|^2}{\langle Ad_k, d_k \rangle} \langle d_k, Ad_k \rangle = 0 \quad (17.488)$$

où nous avons utilisé la valeur (17.486) de  $t_k$ .

- (5) **Calcul de  $f(x_{k+1})$**

Nous repartons de (17.485) où nous substituons la valeur (17.486) de  $t_k$  :

$$f(x_{k+1}) = f(x_k) + \frac{1}{2} \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle} - \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle} = f(x_k) - \frac{1}{2} \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle}. \quad (17.489)$$

- (6) **Encore du calcul ...**

Vu que le produit  $\langle Ad_k, d_k \rangle$  arrive tout le temps, nous allons étudier  $\langle A^{-1}d_k, d_k \rangle$ . Le truc malin est d'essayer d'exprimer ça en termes de  $\bar{x}$  et  $\bar{f} = f(\bar{x})$ . Pour cela nous calculons  $f(\bar{x})$  :

$$\bar{f} = f(\bar{x}) = f(-A^{-1}b) = -\frac{1}{2} \langle b, A^{-1}b \rangle. \quad (17.490)$$

Ayant cela en tête nous pouvons calculer :

$$\langle A^{-1}d_k, d_k \rangle = \langle A^{-1}(Ax_k + b), Ax_k + b \rangle \quad (17.491a)$$

$$= \langle x_k, Ax_k \rangle + \langle A^{-1}b, Ax_k \rangle + \underbrace{\langle b, x_k \rangle + \langle A^{-1}b, b \rangle}_{-2\bar{f}} \quad (17.491b)$$

$$= \langle x_k, Ax_k \rangle + 2\langle x_k, b \rangle - 2\bar{f} \quad (17.491c)$$

$$= 2(f(x_k) - \bar{f}) \quad (17.491d)$$

où nous avons utilisé le fait que  $\langle x, Ay \rangle = \langle Ax, y \rangle$  parce que  $A$  est symétrique.

(7) **Erreur sur la valeur du minimum**

Nous voulons à présent estimer la différence  $f(x_{k+1}) - \bar{f}$ . Pour cela nous mettons en facteur  $f(x_k) - \bar{f}$  dans  $f(x_{k+1}) - \bar{f}$ ; et d'ailleurs c'est pour cela que nous avons calculé  $\langle A^{-1}d_k, d_k \rangle$  : parce que ça fait intervenir  $f(x_k) - \bar{f}$ .

$$f(x_{k+1}) - \bar{f} = f(x_k) - \frac{1}{2} \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle} - \bar{f} \quad (17.492a)$$

$$= (f(x_k) - \bar{f}) \left( 1 - \frac{1}{2} \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle (f(x_k) - \bar{f})} \right) \quad (17.492b)$$

$$= (f(x_k) - \bar{f}) \left( 1 - \frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle \langle A^{-1}d_k, d_k \rangle} \right). \quad (17.492c)$$

Nous traitons le dénominateur à l'aide de l'inégalité de Kantorovitch 17.104. Nous avons

$$\frac{\|d_k\|^4}{\langle Ad_k, d_k \rangle \langle A^{-1}d_k, d_k \rangle} \geq \frac{\|d_k\|^4}{\frac{1}{4} \left( \sqrt{c_2(A)} + \frac{1}{\sqrt{c_2(A)}} \right)^2 \|d_k\|^4} = \frac{4c_2(A)}{(c_2(A) + 1)^2}. \quad (17.493)$$

Mettre cela dans (17.492c) est un calcul d'addition de fractions :

$$f(x_{k+1}) - \bar{f} \leq (f(x_k) - \bar{f}) \left( \frac{c_2(A) - 1}{c_2(A) + 1} \right)^2. \quad (17.494)$$

Par récurrence nous avons alors

$$f(x_k) - \bar{f} \leq (f(x_0) - \bar{f}) \left( \frac{c_2(A) - 1}{c_2(A) + 1} \right)^{2k}. \quad (17.495)$$

Notons qu'il n'y a pas de valeurs absolues parce que  $\bar{f}$  étant le minimum de  $f$ , les deux côtés de l'inégalité sont automatiquement positifs.

(8) **Erreur sur la position du minimum**

Nous voulons à présent étudier la norme de  $x_k - \bar{x}$ . Pour cela nous l'écrivons directement avec la définition de  $f$  en nous souvenant que  $b = -A\bar{x}$  :

$$f(x_k) - \bar{f} = \frac{1}{2} \langle Ax_k, x_k \rangle - \langle A\bar{x}, x_k \rangle + \frac{1}{2} \langle A\bar{x}, \bar{x} \rangle \quad (17.496a)$$

$$= \frac{1}{2} \langle Ax_k, x_k \rangle - \frac{1}{2} \langle A\bar{x}, x_k \rangle - \frac{1}{2} \langle A\bar{x}, x_k \rangle + \frac{1}{2} \langle A\bar{x}, \bar{x} \rangle \quad (17.496b)$$

$$= \frac{1}{2} \left( \langle A(x_k - \bar{x}), x_k \rangle + \langle A\bar{x}, \bar{x} - x_k \rangle \right) \quad (17.496c)$$

$$= \frac{1}{2} \left( \langle A(x_k - \bar{x}), (x_k - \bar{x}) \rangle \right) \quad (17.496d)$$

où à la dernière ligne nous avons fait  $\langle A\bar{x}, \bar{x} - x_k \rangle = \langle \bar{x}, A(\bar{x} - x_k) \rangle$  en vertu de la symétrie de  $A$ .

Les produits de la forme  $\langle Ay, y \rangle$  sont majorés par  $\lambda_{\min} \|y\|^2$  parce que  $\lambda_{\min}$  est la plus petite valeur propre de  $A$ . Dans notre cas,

$$f(x_k) - \bar{f} \geq \frac{1}{2} \lambda_{\min} \|x_k - \bar{x}\|^2 \quad (17.497)$$

(9) **Conclusion**

En combinant les inéquations (17.497) et (17.495) nous trouvons

$$\frac{1}{2}\lambda_{\min}\|x_k - \bar{x}\|^2 \leq f(x_k) - \bar{f} \leq (f(x_0) - \bar{f}) \left( \frac{c_2(A) - 1}{c_2(A) + 1} \right)^{2k}, \quad (17.498)$$

c'est-à-dire

$$\|x_k - \bar{x}\| \leq \sqrt{\frac{2(f(x_0) - \bar{f})}{\lambda_{\min} + 1}}^{2k}. \quad (17.499)$$

□

Notons que lorsque  $c_2(A)$  est proche de 1 la méthode converge rapidement. Par contre si  $c_2(A)$  est proche de zéro, la méthode converge lentement.

## 17.14 Formes quadratiques, signature, et lemme de Morse

### 17.115.

NORMooHSWKooLtUbRl

Soit  $(E, \|\cdot\|_E)$  un espace vectoriel réel normé de dimension finie  $n$ . L'ensemble des formes quadratiques réelles<sup>62</sup> sur  $E$  est vu comme l'ensemble des matrices symétriques  $S_n(\mathbb{R})$ ; il sera noté  $Q(E)$  et le sous-ensemble des formes quadratiques non dégénérées est  $S_n(\mathbb{R}) \cap \text{GL}(n, \mathbb{R})$  qui sera noté  $\Omega(E)$ . Nous rappelons que la correspondance est donnée de la façon suivante.

Si  $A \in S_n(\mathbb{R})$ , la forme quadratique associée est  $q_A$  donnée par le produit scalaire  $q_A(x) = x \cdot Ax$ .

Pour information, le lemme de Morse est le lemme 20.198.

NORMooQZFLooYnILtn

### 17.116.

Nous noterons encore  $Q^+(E)$  les formes quadratiques positives sur  $E$  et  $Q^{++}(E)$  les formes quadratiques strictement définies positives sur  $E$ .

Sur  $Q(E)$  nous mettons la norme

$$N(q) = \sup_{\|x\|_E=1} |q(x)|, \quad (17.500)$$

qui du point de vue de  $S_n(\mathbb{R})$  est

$$N(A) = \sup_{\|x\|_E=1} |x^t Ax|. \quad (17.501)$$

EqDQgBNAg

Notons que à droite, c'est la valeur absolue usuelle sur  $\mathbb{R}$ .

Nous savons par le théorème de Sylvester (théorème 8.250) que dans  $M(n, \mathbb{R})$ , toute matrice symétrique de signature  $(p, q)$  est semblable à la matrice

$$\mathbb{I}_{p,q} = \begin{pmatrix} -\mathbb{I}_q & & \\ & \mathbb{I}_p & \\ & & 0_{n-p-q} \end{pmatrix}. \quad (17.502)$$

Donc deux matrices de  $S_n$  sont semblables si et seulement si elles ont la même signature (même si elles ne sont pas de rang maximum, cela soit dit au passage). Si nous notons  $S_n^{p,q}(\mathbb{R})$  l'ensemble des matrices réelles symétriques de signature  $(p, q)$ , alors

$$S_n^{p,q}(\mathbb{R}) = \{P^t AP \text{ tel que } P \in \text{GL}(n, \mathbb{R})\} \quad (17.503)$$

où  $A$  est une quelconque de ces matrices.

---

62. Définition 8.122.

Nous voudrions en savoir plus sur ces ensembles. En particulier nous aimerions savoir si la signature est une notion « stable » au sens où ces ensembles seraient ouverts dans  $S_n$ . Pour cela nous considérons l'action de  $\mathrm{GL}(n, \mathbb{R})$  sur  $S_n$  définie par

$$\begin{aligned}\alpha: \mathrm{GL}(n, \mathbb{R}) \times S_n(\mathbb{R}) &\rightarrow S_n(\mathbb{R}) \\ (P, A) &\mapsto P^t A P\end{aligned}\tag{17.504}$$

faite exprès pour que les orbites de cette action soient les ensembles  $S_n^{p,q}(\mathbb{R})$ .

La proposition suivante montre que lorsque  $p + q = n$ , c'est-à-dire lorsqu'on parle de matrices de rang maximum, les ensembles  $S_n^{p,q}(\mathbb{R})$  sont ouverts, c'est-à-dire que la signature d'une forme quadratique est une propriété « stable » par petite variations des éléments de matrice. Notons tout de suite que si le rang n'est pas maximum, le théorème de Sylvester dit qu'elle est semblable à une matrice diagonale avec des zéros sur la diagonale ; en modifiant un peu ces zéros, on peut modifier évidemment la signature.

PropNPbnsMd

**Proposition 17.117** ([126]).

Soit  $(E, \|\cdot\|_E)$  un espace vectoriel normé de dimension finie. Alors

- (1) les formes quadratiques non dégénérées forment un ouvert dans l'ensemble des formes quadratiques,
- (2) les ensembles  $S_n^{p,q}(\mathbb{R})$  avec  $p + q = n$  sont ouverts dans  $S_n(\mathbb{R})$ ,
- (3) les composantes connexes de  $\Omega(E)$  sont les  $S_n^{p,q}(\mathbb{R})$  avec  $p + q = n$ ,
- (4) les  $S_n^{p,q}(\mathbb{R})$  non dégénérés sont connexes par arc.

ItemGOhRliViii

ItemGOhRliViv

*Démonstration.* Cette preuve est donnée du point de vue des matrices. La différence entre le point (3) et (4) est que dans le premier nous prouvons la connexité de  $S_n^{p,q}(\mathbb{R})$  à partir de la connexité de  $\mathrm{GL}^+(n, \mathbb{R})$ , tandis que dans le second nous prouvons la connexité par arc de  $S_n^{p,q}(\mathbb{R})$  à partir de la connexité par arc de  $\mathrm{GL}^+(n, \mathbb{R})$ . Bien entendu le second implique le premier.

- (1) Il s'agit simplement de remarquer que  $Q(E) = S_n(\mathbb{R})$ , que  $\Omega(E) = S_n(\mathbb{R}) \cap \mathrm{GL}(n, \mathbb{R})$  et que le déterminant est une fonction continue sur  $\mathrm{M}(n, \mathbb{R})$ .
- (2) Soit  $A_0 \in S_n^{p,q}(\mathbb{R})$ . Le théorème de Sylvester 8.250 nous donne une matrice inversible  $P$  telle que  $P^t A_0 P = \mathbb{1}_{p,q}$ . Nous allons montrer qu'il existe un voisinage  $\mathcal{U}$  de  $\mathbb{1}_{p,q}$  contenu dans  $S_n^{p,q}(\mathbb{R})$ . À partir de là, l'ensemble  $(P^{-1})^t \mathcal{U} P^{-1}$  sera un voisinage de  $A_0$  contenu dans  $S_n^{p,q}(\mathbb{R})$ .

Nous considérons les espaces vectoriels

$$F = \mathrm{Span}\{e_1, \dots, e_p\} \tag{17.505a}$$

$$G = \mathrm{Span}\{e_{p+1}, \dots, e_n\} \tag{17.505b}$$

La norme euclidienne  $\|\cdot\|_p$  sur  $F$  est équivalente à la norme  $\|\cdot\|_E$  par le théorème 10.57. Donc il existe une constante  $k_1 > 0$  telle que pour tout  $x \in F$ ,

$$\|x\|_p \geq k_1 \|x\|_E. \tag{17.506}$$

EqMVjCijJ

De la même façon sur  $G$ , il existe une constante  $k_2 > 0$  telle que

$$\|x\|_q \geq k_2 \|x\|_E. \tag{17.507}$$

EqSEwOcDw

Si nous posons  $k = \min\{k_1^2, k_2^2\}$ , alors nous avons

$$\forall x \in F, \quad \|x\|_p^2 \geq k_1^2 \|x\|_E^2 \geq k \|x\|_E^2 \tag{17.508a}$$

$$\forall x \in G, \quad \|x\|_q^2 \geq k_2^2 \|x\|_E^2 \geq k \|x\|_E^2. \tag{17.508b}$$

Soit une matrice  $A \in S_n(\mathbb{R})$  telle que  $N(A - \mathbb{1}_{p,q}) < k$ , c'est-à-dire que  $A$  est dans un voisinage de  $\mathbb{1}_{p,q}$  pour la norme sur  $S_n(\mathbb{R})$  donné par (17.501). Si  $x$  est non nul dans  $E$ , nous avons

$$|x^t (A - \mathbb{1}_{p,q}) x| \leq N(\mathbb{1}_{p,q} - A) \|x\|^2 \leq k \|x\|^2. \tag{17.509}$$

En déballant la valeur absolue, cela signifie que

$$-k\|x\|_E^2 \leq x^t(A - \mathbb{1}_{p,q})x \leq k\|x\|^2. \quad (17.510)$$

Si  $x \in F$ , alors la première inéquation et (17.506) donnent

$$x^t Ax \geq \|x\|_p^2 - k\|x\|_E^2 > 0 \quad (17.511)$$

Si  $x \in G$ , alors la seconde inéquation et (17.507) donnent

$$x^t Ax \leq k\|x\|_E^2 - \|x\|_q^2 < 0. \quad (17.512)$$

Nous avons donc montré que  $x \mapsto x^t Ax$  est positive sur  $F$  et négative sur  $G$ , ce qui prouve que  $A$  est bien de signature  $(p, q)$  et appartient donc à  $S_n^{p,q}(\mathbb{R})$ . Autrement dit nous avons

$$B(\mathbb{1}_{p,q}, k) \subset S_n^{p,q}(\mathbb{R}). \quad (17.513)$$

- (3) Cette partie de la preuve provient essentiellement de [586], et fonctionne pour tous les  $S_n^{p,q}(\mathbb{R})$ , même pour ceux qui ne sont pas de rang maximum.

Soit  $A \in S_n^{p,q}(\mathbb{R})$ . Nous savons que  $\mathrm{GL}(n, \mathbb{R})$  contient exactement deux composantes connexes (proposition 13.26). Vu que l'application

$$\begin{aligned} \alpha: \mathrm{GL}(n, \mathbb{R}) &\rightarrow S_n \\ P &\mapsto P^t AP \end{aligned} \quad (17.514)$$

est continue, l'image d'un connexe de  $\mathrm{GL}(n, \mathbb{R})$  par  $\alpha$  est connexe (lemme 7.272). En particulier,  $\alpha(\mathrm{GL}^\pm(n, \mathbb{R}))$  sont deux connexes et nous savons que  $S_n^{p,q}(\mathbb{R})$  a au plus ces deux composantes connexes.

Notre but est maintenant de trouver une intersection entre les parties  $\alpha(\mathrm{GL}^+(n, \mathbb{R}))$  et  $\alpha(\mathrm{GL}^-(n, \mathbb{R}))$ <sup>63</sup>. Soit par le théorème de Sylvester, soit par le théorème de diagonalisation des matrices symétriques réelles 8.226, il existe une matrice  $P \in \mathrm{GL}(n, \mathbb{R})$  diagonalisant  $A$ . En suivant la remarque 8.227, et en notant  $Q$  la matrice obtenue à partir de  $P$  en changeant le signe de sa première ligne, nous avons

$$\alpha(Q) = Q^t A Q = P^t A P = \alpha(P). \quad (17.515)$$

Or si  $P \in \mathrm{GL}^+(n, \mathbb{R})$ , alors  $Q \in \mathrm{GL}^-(n, \mathbb{R})$  et inversement. Donc nous avons trouvé une intersection entre  $\alpha(\mathrm{GL}^+(n, \mathbb{R}))$  et  $\alpha(\mathrm{GL}^-(n, \mathbb{R}))$ .

- (4) Soient  $A$  et  $B$  dans  $S_n^{p,q}(\mathbb{R}) \cap \mathrm{GL}(n, \mathbb{R})$ . Par le théorème de Sylvester, il existe  $P$  et  $Q$  dans  $\mathrm{GL}(n, \mathbb{R})$  telles que  $A = P^t \mathbb{1}_{p,q} P$  et  $B = Q^t \mathbb{1}_{p,q} Q$ . Par la remarque 8.227 nous pouvons choisir  $P$  et  $Q$  dans  $\mathrm{GL}^+(n, \mathbb{R})$ . Ce dernier groupe étant connexe par arc, il existe un chemin

$$\gamma: [0, 1] \rightarrow \mathrm{GL}^+(n, \mathbb{R}) \quad (17.516)$$

tel que  $\gamma(0) = P$  et  $\gamma(1) = Q$ . Alors le chemin

$$s \mapsto \gamma(s)^t \mathbb{1}_{p,q} \gamma(s) \quad (17.517)$$

est un chemin continu dans  $S_n^{p,q}(\mathbb{R})$  joignant  $A$  à  $B$ .

□

Nous savons déjà de la proposition 17.117 que les ensembles  $S_n^{p,q}(\mathbb{R})$  (pas spécialement de rang maximum) sont ouverts dans  $S_n(\mathbb{R})$ . Le lemme suivant nous donne une précision à ce sujet, dans le cas des matrices de rang maximum, en disant que la matrice qui donne la similitude entre  $A_0$  et  $A$  est localement un  $C^1$ -difféomorphisme de  $A$ .

63. À ce point, il me semble que [586] fait erreur parce que la matrice  $-\mathbb{1}_n$  est de déterminant 1 lorsque  $n$  est pair. L'argument donné ici provient de [126]

LemWLCvLXe

**Lemme 17.118.**

Soit  $A_0 \in \Omega(\mathbb{R}^n) = S_n \cap \text{GL}(n, \mathbb{R})$ , une matrice symétrique inversible. Alors il existe un voisinage  $V$  de  $A_0$  dans  $S_n$  et une application  $\varphi: V \rightarrow \text{GL}(n, \mathbb{R})$  qui

- (1) est de classe  $C^1$ ,
- (2) est telle que pour tout  $A \in V$ ,  $\varphi(A)^t A_0 \varphi(A) = A$ .

*Démonstration.* Nous considérons l'application

$$\begin{aligned} \varphi: \mathbb{M}(n, \mathbb{R}) &\rightarrow S_n \\ M &\mapsto M^t A_0 M. \end{aligned} \quad (17.518)$$

Étant donné que les composantes de  $\varphi(M)$  sont des polynômes en les entrées de  $M$ , cette application est de classe  $C^1$  – et même plus. Soit maintenant  $H \in \mathbb{M}(n, \mathbb{R})$  et calculons  $d\varphi_{\mathbb{1}}(H)$  par la formule (12.715) :

$$d\varphi_{\mathbb{1}}(H) = \frac{d}{dt} \left[ \varphi(\mathbb{1} + tH) \right]_{t=0} \quad (17.519a)$$

$$= \frac{d}{dt} \left[ (\mathbb{1} + tH^t) A_0 (\mathbb{1} + tH) \right]_{t=0} \quad (17.519b)$$

$$= \frac{d}{dt} \left[ A_0 + tA_0 H + tH^t A_0 + t^2 H^t A_0 H \right]_{t=0} \quad (17.519c)$$

$$= A_0 H + H^t A_0. \quad (17.519d)$$

Donc

$$d\varphi_{\mathbb{1}}(H) = (A_0 H) + (A_0 H)^t. \quad (17.520)$$

Par conséquent

$$\ker(d\varphi_{\mathbb{1}}) = \{H \in \mathbb{M}(n, \mathbb{R}) \text{ tel que } A_0 H \text{ est antisymétrique}\}, \quad (17.521)$$

et si nous posons

$$F = \{H \in \mathbb{M}(n, \mathbb{R}) \text{ tel que } A_0 H \text{ est symétrique}\} \quad (17.522)$$

nous avons

$$\mathbb{M}(n, \mathbb{R}) = F \oplus \ker(d\varphi_{\mathbb{1}}) \quad (17.523)$$

parce que toute matrice peut être décomposée de façon unique en partie symétrique et antisymétrique. De plus l'application

$$\begin{aligned} f: F &\rightarrow S_n \\ H &\mapsto A_0 H \end{aligned} \quad (17.524) \quad \text{EqGTBpsDm}$$

est une bijection linéaire. D'abord  $A_0 H = 0$  implique  $H = 0$  parce que  $A_0$  est inversible, et ensuite si  $X \in S_n$ , alors  $X = A_0 A_0^{-1} X$ , ce qui prouve que  $X$  est l'image par  $f$  de  $A_0^{-1} X$  et donc que  $f$  est surjective.

Maintenant nous considérons la restriction  $\psi = \varphi|_F$ ,  $\psi: F \rightarrow S_n$ . Remarquons que  $\mathbb{1} \in F$  parce que  $A_0 \in S_n$ . L'application  $d\psi_{\mathbb{1}}$  est une bijection. En effet d'abord

$$d(\varphi|_F)_{\mathbb{1}} = (d\varphi_{\mathbb{1}})|_F, \quad (17.525)$$

ce qui prouve que

$$\ker(d\psi_{\mathbb{1}}) = \ker(d\varphi_{\mathbb{1}}) \cap F = \{0\}, \quad (17.526)$$

ce qui prouve que  $d\psi_{\mathbb{1}}$  est injective. Pour montrer que  $d\psi_{\mathbb{1}}$  est surjective, il suffit de mentionner le fait que  $\dim F = \dim S_n$  du fait que l'application (17.524) est une bijection linéaire.

Nous pouvons utiliser le théorème d'inversion locale (théorème 17.47) et conclure qu'il existe un voisinage ouvert  $U$  de  $\mathbb{1}$  dans  $F$  tel que  $\psi$  soit un difféomorphisme  $C^1$  entre  $U$  et  $V = \psi(U)$ .



Vu que  $\mathrm{GL}(n, \mathbb{R})$  est ouvert dans  $\mathrm{M}(n, \mathbb{R})$ , nous pouvons prendre  $U \cap \mathrm{GL}(n, \mathbb{R})$  et donc supposer que  $U \subset \mathrm{GL}(n, \mathbb{R})$ .

Pour tout  $A \in V$ , il existe une unique  $M \in U$  telle que  $\psi(M) = A$ , c'est-à-dire telle que  $A = M^t A_0 M$ . Cette matrice  $M$  est  $\psi^{-1}(A)$  et est une matrice inversible. Bref, nous posons

$$\begin{aligned}\phi: V &\rightarrow \mathrm{GL}(n, \mathbb{R}) \\ A &\mapsto \psi^{-1}(A),\end{aligned}\tag{17.527}$$

et ce  $\phi$  est de classe  $C^1$  sur  $V$  parce que c'est ce que dit le théorème d'inversion locale. Cette application répond à la question parce que  $V$  est un voisinage de  $\varphi(\mathbb{1}) = A_0$  et pour tout  $A \in V$  nous avons

$$\phi(A)^t A_0 \phi(A) = \varphi^{-1}(A)^t A_0 \varphi^{-1}(A) = A.\tag{17.528}$$

□

## 17.15 Ellipsoïde de John-Loewner

C'est le moment de relire les conventions de notations données en 17.115 et 17.116 ainsi que la définition 8.122 d'une forme quadratique.

Soit  $q$  une forme quadratique sur  $\mathbb{R}^n$  ainsi que  $\mathcal{B}$  une base orthonormée de  $\mathbb{R}^n$  dans laquelle la matrice de  $q$  est diagonale. Dans cette base, la forme  $q$  est donnée par la proposition 8.254 :

$$q(x) = \sum_i \lambda_i x_i^2\tag{17.529}$$

où les  $\lambda_i$  sont les valeurs propres de  $q$ .

Plus généralement nous notons  $\mathrm{mat}_{\mathcal{B}}(q)$  la matrice de  $q$  dans la base  $\mathcal{B}$  de  $\mathbb{R}^n$ .

PropOXWooYrDKpw

### Proposition 17.119.

Soit  $\mathcal{B}$  une base orthonormée de  $\mathbb{R}^n$  et l'application<sup>64</sup>

$$\begin{aligned}D: Q(\mathbb{R}^n) &\rightarrow \mathbb{R} \\ q &\mapsto \det(\mathrm{mat}_{\mathcal{B}}(q)).\end{aligned}\tag{17.530}$$

Alors :

- (1) La valeur de  $D$  ne dépend pas du choix de la base orthonormée  $\mathcal{B}$ .
- (2) La fonction  $D$  est donnée par la formule  $D(q) = \prod_i \lambda_i$  où les  $\lambda_i$  sont les valeurs propres de  $q$ .
- (3) La fonction  $D$  est continue.

*Démonstration.* Soit  $q$  une forme quadratique sur  $\mathbb{R}^n$ . Nous considérons  $\mathcal{B}$  une base de diagonalisation de  $q$  :

$$q(x) = \sum_i \lambda_i x_i^2\tag{17.531}$$

où les  $x_i$  sont les composantes de  $x$  dans la base  $\mathcal{B}$ . Par définition, la matrice  $\mathrm{mat}_{\mathcal{B}}(q)$  est la matrice diagonale contenant les valeurs propres de  $q$ .

Nous considérons aussi  $\mathcal{B}_1$ , une autre base orthonormée de  $\mathbb{R}^n$ . Nous notons  $S = \mathrm{mat}_{\mathcal{B}_1}(q)$  ; étant symétrique, cette matrice se diagonalise par une matrice orthogonale : il existe  $P \in \mathrm{O}(n, \mathbb{R})$  telle que

$$S = P \mathrm{mat}_{\mathcal{B}}(q) P^t;\tag{17.532}$$

donc  $\det(S) = \det(P P^t) \det(\mathrm{diag}(\lambda_1, \dots, \lambda_n)) = \lambda_1 \dots \lambda_n$ . Ceci prouve en même temps que  $D$  ne dépend pas du choix de la base et que sa valeur est le produit des valeurs propres.

Passons à la continuité. L'application déterminant  $\det: S_n(\mathbb{R}^n) \rightarrow \mathbb{R}$  est continue car polynôme en les composantes. D'autre part l'application  $\mathrm{mat}_{\mathcal{B}}: Q(\mathbb{R}^n) \rightarrow S_n(\mathbb{R})$  est continue par la proposition 8.139. L'application  $D$  étant la composée de deux applications continues, elle est continue. □

64. L'ensemble  $Q(E)$  est l'ensemble des formes quadratiques sur  $E$ .

LEMooLSTOooZiEOdx

**Lemme 17.120** (Volume d'un ellipsoïde[126, 1]).

Soit une forme quadratique strictement définie positive  $q$  sur  $\mathbb{R}^n$ . Nous considérons l'ellipsoïde

$$\mathcal{E} = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leq r^2\} \quad (17.533)$$

pour un certain  $r > 0$ .

Alors le volume de  $\mathcal{E}$  est donné par

$$V_q = \frac{V_0 r^n}{\sqrt{D(q)}} \quad (17.534)$$

où  $V_0$  est une constante<sup>65</sup> et  $D$  est l'application de la proposition 17.119.

*Démonstration.* Vu que la matrice de  $q$  est symétrique, elle est diagonalisable par une matrice orthogonale (théorème 8.226). Autrement dit, il existe une base orthonormée  $\mathcal{B} = \{e_1, \dots, e_n\}$  de  $\mathbb{R}^n$  telle que

$$q(x) = \sum_{i=1}^n a_i x_i^2 \quad (17.535)$$

EqELBooQLPQUi

où  $x_i = \langle e_i, x \rangle$  et les  $a_i$  sont tous strictement positifs.

À priori nous devrions calculer

$$\int_{\mathcal{E}} dx, \quad (17.536)$$

mais nous effectuons le changement de variable<sup>66</sup> associé à la matrice qui diagonalise  $q$  et nous devons simplement calculer

$$V_q = \int_{\sum_i a_i x_i^2 < r^2} dx \quad (17.537)$$

parce que le jacobien de ce changement de variable est 1 (déterminant d'une matrice orthogonale).

Tout cela pour dire que nous nommons  $\mathcal{E}_q$  l'ellipsoïde associée à la forme quadratique  $q$  et  $V_q$  son volume que nous allons maintenant calculer<sup>67</sup> :

$$V_q = \int_{\sum_i a_i x_i^2 < r^2} dx \quad (17.538)$$

Cette intégrale est écrite de façon plus simple en utilisant le  $C^1$ -difféomorphisme

$$\begin{aligned} \varphi: \mathcal{E}_q &\rightarrow B(0, 1) \\ x &\mapsto \frac{1}{r} (x_1 \sqrt{a_1}, \dots, x_n \sqrt{a_n}). \end{aligned} \quad (17.539)$$

Le fait que  $\varphi$  prenne bien ses valeurs dans  $B(0, 1)$  est un simple calcul : si  $x \in \mathcal{E}_q$ , alors

$$\sum_i \varphi(x)_i^2 = \frac{1}{r^2} \sum_i a_i x_i^2 = \frac{1}{r^2} q(x) < 1. \quad (17.540)$$

Cela nous permet d'utiliser le théorème de changement de variables 14.282 :

$$V_q = \int_{\sum_i a_i x_i^2 < r^2} dx = \frac{r^n}{\sqrt{a_1 \dots a_n}} \int_{B(0,1)} dx. \quad (17.541)$$

La dernière intégrale est le volume de la boule unité dans  $\mathbb{R}^n$  et nous la notons  $V_0$ . La proposition 17.119 nous permet d'écrire  $V_q$  sous la forme

$$V_q = \frac{V_0 r^n}{\sqrt{D(q)}}. \quad (17.542)$$

□

65. C'est le volume de la boule unité dans  $\mathbb{R}^n$  et ce n'est pas tout à fait évident à calculer [587].

66. Théorème 14.282.

67. Le volume ne change pas si nous écrivons l'inégalité stricte au lieu de large dans le domaine d'intégration ; nous le faisons pour avoir un domaine ouvert.

**17.121.**

Le théorème suivant dit en substance que si  $K$  est compact, alors il existe un unique ellipsoïde de volume minimal centré en l'origine et contenant  $K$ . Il faut se rendre compte que l'ellipsoïde n'est pas celui que l'on croirait intuitivement parce que la contrainte *centrée en l'origine* est forte. Si  $K = \overline{B((4,0),1)}$ , alors l'ellipsoïde donnée n'est pas du tout  $K$  lui-même comme on pourrait s'y attendre. Ce serait probablement quelque chose comme la boule centrée en  $(0,0)$  et de rayon 5<sup>68</sup>.

Ce que l'on voudrait est un ellipsoïde qui soit centré où il faut pour que le volume soit minimal. Nous verrons que c'est possible en la proposition 17.123, mais qu'alors l'unicité est moins évidente (voir la remarque dans [588]). Si vous voulez en entendre parler, vous pouvez lire [589, 590].

PropJYVooRMaPok

**Proposition 17.122** (Ellipsoïde de John-Loewner[126]).

Soit  $K$  compact dans  $\mathbb{R}^n$  et d'intérieur non vide. Il existe une unique ellipsoïde<sup>69</sup> (pleine) centrée en l'origine de volume minimal contenant  $K$ .

*Démonstration.* Nous subdivisons la preuve en plusieurs parties.

- (1) **Volume de l'ellipsoïde** Soit  $\mathcal{E}$  un ellipsoïde centré en l'origine. La proposition 8.315 et son corolaire 8.316 nous indiquent que

$$\mathcal{E} = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leq 1\} \quad (17.543)$$

pour une certaine forme quadratique strictement définie positive  $q$ . Le lemme 17.120 nous donne alors le volume de  $\mathcal{E}$  par

$$V_q = \frac{V_0}{\sqrt{D(q)}} \quad (17.544)$$

où  $V_0$  est une constante.

- (2) **Existence de l'ellipsoïde**

Nous voulons trouver un ellipsoïde contenant  $K$  de volume minimal, c'est-à-dire une forme quadratique  $q \in Q^{++}(\mathbb{R}^n)$  telle que

- $D(q)$  soit maximal
- $q(x) \leq 1$  pour tout  $x \in K$ .

Nous considérons l'ensemble des candidats semi-définis positifs.

$$A = \{q \in Q^+ \text{ tel que } q(x) \leq 1 \forall x \in K\}. \quad (17.545)$$

Nous allons montrer que  $A$  est convexe, compact et non vide dans  $Q(\mathbb{R}^n)$ ; il aura ainsi un maximum de la fonction continue  $D$  définie sur  $Q(\mathbb{R}^n)$ . Nous montrerons ensuite que le maximum est dans  $Q^{++}$ . L'unicité sera prouvée à part.

- (2a) **Non vide** L'ensemble  $K$  est compact et donc borné par  $M > 0$ . La forme quadratique  $q: x \mapsto \|x\|^2/M^2$  est dans  $A$  parce que si  $x \in K$  alors

$$q(x) = \frac{\|x\|^2}{M^2} \leq 1. \quad (17.546)$$

- (2b) **Convexe** Soient  $q, q' \in A$  et  $\lambda \in [0, 1]$ . Nous avons encore  $\lambda q + (1 - \lambda)q' \in Q^+$  parce que

$$\lambda q(x) + (1 - \lambda)q'(x) \geq 0 \quad (17.547)$$

dès que  $q(x) \geq 0$  et  $q'(x) \geq 0$ . D'autre part si  $x \in K$  nous avons

$$\lambda q(x) + (1 - \lambda)q'(x) \leq \lambda + (1 - \lambda) = 1. \quad (17.548)$$

Donc  $\lambda q + (1 - \lambda)q' \in A$ .

68. Je ne sais pas très bien si il y a moyen de faire mieux. Ce serait sans doute un bon exercice ; faites-moi savoir si vous avez la réponse.

69. Définition 8.313.

(2c) **Fermé**

Pour rappel, la topologie de  $Q(\mathbb{R}^n)$  est celle de la norme (8.293). Nous considérons une suite  $(q_n)$  dans  $A$  convergeant vers  $q \in Q(\mathbb{R}^n)$  et nous allons prouver que  $q \in A$ , de sorte que la caractérisation séquentielle de la fermeture (proposition 7.154) conclue que  $A$  est fermé. En nommant  $e_x$  le vecteur unitaire dans la direction  $x$  nous avons

$$|q(x)| = \|\|x\|^2 q(e_x)\| \leq \|x\|^2 N(q), \quad (17.549)$$

de sorte que notre histoire de suite convergente donne pour tout  $x$  :

$$|q_n(x) - q(x)| \leq \|x\|^2 N(q_n - q) \rightarrow 0. \quad (17.550)$$

Vu que  $q_n(x) \geq 0$  pour tout  $n$ , nous devons aussi avoir  $q(x) \geq 0$  et donc  $q \in Q^+$  (semi-définie positive). De la même manière si  $x \in K$  alors  $q_n(x) \leq 1$  pour tout  $n$  et donc  $q(x) \leq 1$ . Par conséquent  $q \in A$  et  $A$  est fermé.

(2d) **Borné**

La partie  $K$  de  $\mathbb{R}^n$  est borné et d'intérieur non vide, donc il existe  $a \in K$  et  $r > 0$  tel que  $\overline{B(a, r)} \subset K$ . Si par ailleurs  $q \in A$  et  $x \in \overline{B(0, r)}$  nous avons  $a + x \in K$  et donc  $q(a + x) \leq 1$ . De plus  $q(-a) = q(a) \leq 1$ , donc

$$\sqrt{q(x)} = \sqrt{q(x + a - a)} \leq \sqrt{q(x + a)} + \sqrt{q(-a)} \leq 2 \quad (17.551)$$

par l'inégalité de Minkowski 10.20. Cela prouve que si  $x \in \overline{B(0, r)}$  alors  $q(x) \leq 4$ . Si par contre  $x \in \overline{B(0, 1)}$  alors  $rx \in \overline{B(0, r)}$  et

$$0 \leq q(x) = \frac{1}{r^2} q(rx) \leq \frac{4}{r^2}, \quad (17.552)$$

ce qui prouve que  $N(q) \leq \frac{4}{r^2}$  et que  $A$  est borné.

L'ensemble  $A$  est compact parce que fermé et borné, théorème de Borel-Lebesgue 9.24. L'application continue  $D: Q(\mathbb{R}^n) \rightarrow \mathbb{R}$  de la proposition 17.119 admet donc un maximum sur le compact  $A$ . Soit  $q_0$  ce maximum.

Nous montrons que  $q_0 \in Q^{++}(\mathbb{R}^d)$ . Nous savons que l'application  $f: x \mapsto \frac{\|x\|^2}{M^2}$  est dans  $A$  et que  $D(f) > 0$ . Vu que  $q_0$  est maximale pour  $D$ , nous avons

$$D(q_0) \geq D(f) > 0. \quad (17.553)$$

Donc  $q_0 \in Q^{++}$ .

(3) **Unicité**

Si il existe une autre ellipsoïde de même volume que celle associée à la forme quadratique  $q_0$ , nous avons une forme quadratique  $q \in Q^{++}$  telle que  $q(x) \leq 1$  pour tout  $x \in K$ . C'est-à-dire que nous avons  $q_0, q \in A$  tels que  $D(q_0) = D(q)$ .

Nous considérons la base canonique  $\mathcal{B}_c$  de  $\mathbb{R}^n$  et nous posons  $S = \text{mat}_{\mathcal{B}_c}(q)$ ,  $S_0 = \text{mat}_{\mathcal{B}_c}(q_0)$ . Étant donné que  $A$  est convexe,  $(q_0 + q)/2 \in A$  et nous allons prouver que cet élément de  $A$  contredit la maximalité de  $q_0$ . En effet

$$D\left(\frac{q + q_0}{2}\right) = \det\left(\frac{S + S_0}{2}\right) \quad (17.554)$$

Nous allons utiliser le lemme 17.92 qui dit que le déterminant est log-concave sous la forme de l'équation (17.361) avec  $\alpha = \beta = \frac{1}{2}$  :

$$D\left(\frac{q + q_0}{2}\right) = \det\left(\frac{S + S_0}{2}\right) > \sqrt{\det(S)}\sqrt{\det(S_0)} = \det(S_0) = D(q_0). \quad (17.555)$$

Nous avons utilisé le fait que  $D(q_0) = D(q)$  qui signifie que  $\det(S_0) = \det(S)$ . L'inéquation (17.555) contredit la maximalité de  $D(q_0)$  et donne donc l'unicité.

□

Dans la proposition suivante nous oublions l'unicité, mais nous démontrons qu'il existe un ellipsoïde de volume minimal parmi les ellipsoïdes centrées où l'on veut et non seulement en zéro. La source de cette proposition est [1], et comme toujours avec cette source, vous devez regarder à la fois l'énoncé et la preuve avec un oeil encore plus prudent que d'habitude.

PROPooVIDPooOGrRJh

**Proposition 17.123** ([1]).

Soit un compact d'intérieur non vide  $K$  dans  $\mathbb{R}^n$ . Il existe un ellipsoïde de volume minimal contenant  $K$ .

*Démonstration.* Au lieu de considérer le compact  $K$  et de chercher où centrer l'ellipsoïde afin qu'elle puisse contenir  $K$  en un volume minimal, nous allons chercher comment translater  $K$  pour qu'un ellipsoïde centré en zéro puisse contenir l'image translétée de  $K$  en un volume minimal.

Pour  $a \in \mathbb{R}^n$ , nous notons  $\mathcal{E}_a$  la plus petite ellipsoïde centrée en 0 contenant  $K + a$  (translation de  $K$  par le vecteur  $a$ ). Elle est bien définie par la proposition 17.122. Notre jeu est maintenant d'étudier la fonction<sup>70</sup>

$$\begin{aligned} f: \mathbb{R}^n &\rightarrow \mathbb{R}^+ \\ a &\mapsto \text{vol}(\mathcal{E}_a). \end{aligned} \quad (17.556)$$

Nous allons montrer que  $f$  est continue et qu'elle peut être restreinte à un compact sans risque de rater un minimum.

- (1) **À propos de forme quadratique** Nous notons  $q$  la forme quadratique de l'ellipsoïde  $\mathcal{E}_0$  et  $A$  sa matrice symétrique associée. Si  $x \in K + h$  nous avons, pour un certain  $k \in K$  :

$$q(x) = q(k + h) \quad (17.557a)$$

$$= \langle k + h, A(k + h) \rangle \quad (17.557b)$$

$$= \langle k, Ak \rangle + \langle k, Ah \rangle + \langle h, Ak \rangle + \langle h, Ah \rangle. \quad (17.557c)$$

En tenant compte du fait que  $A = A^t$  nous avons aussi  $\langle k, Ah \rangle = \langle h, Ak \rangle$  et donc

$$q(x) = 2\langle k, Ah \rangle + q(h). \quad (17.558)$$

En ce qui concerne la norme, pour tout  $x \in K + h$  nous avons donc la majoration

$$|q(x)| \leq 1 + 2\|k\|\|Ah\| + q(h) \quad (17.559a)$$

$$\leq 1 + 2|K|\|Ah\| + q(h) \quad (17.559b)$$

où nous avons noté  $|K| = \sup_{k \in K} \|k\|$ . Vu que  $q(x)$  est en fait toujours positif nous pouvons oublier la valeur absolue à gauche et conclure que  $K + h$  est contenu dans l'ellipsoïde

$$\mathcal{F}_1 = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leq 1 + 2|K|\|Ah\| + q(h)\}. \quad (17.560)$$

(2) **Volumes**

Nous avons donc  $\text{vol}(\mathcal{E}_h) \leq \text{vol}(\mathcal{F}_1)$ . En reprenant la formule du lemme 17.120 pour le volume de l'ellipsoïde nous avons la majoration

$$\text{vol}(\mathcal{E}_h) \leq \text{vol}(\mathcal{E}_0)(1 + 2|K|\|Ah\| + q(h))^n. \quad (17.561)$$

Autrement dit, en ce qui concerne notre fonction  $f$  :

$$f(h) \leq f(0)(1 + 2|K|\|Ah\| + q(h))^n. \quad (17.562)$$

EQooLVZCooIVxVNX

Nous pouvons faire le même raisonnement en partant de  $K + h$  et en voyant comment modifier le rayon de  $\mathcal{E}_h$  pour que  $\mathcal{E}_0$  y rentre. Autrement dit, nous refaisons le raisonnement en posant

<sup>70</sup>. Notons que cette fonction est bien définie, même sans la partie unicité de 17.122 parce que même si  $\mathcal{E}_a$  n'était pas bien définie, son volume, lui, est bien défini.

$K' = K + h$  et en étudiant  $K' - h$ . Si  $q_1$  est la forme quadratique de  $\mathcal{E}_h$  et si  $A_1$  est sa matrice symétrique,

$$\text{vol}(\mathcal{E}_0) \leq \text{vol}(\mathcal{E}_h) (1 + 2\|K + h\| \|A_1 h\| + q_1(h))^n \quad (17.563)$$

En termes de  $f$  :

$$f(0) \leq f(h) (1 + 2\|K + h\| \|A_1 h\| + q_1(h))^n \quad (17.564)$$

### (3) Encadrement

Les majoration (17.562) et (17.564) nous permettent de créer un encadrement pour  $f(h)$ . En effet, en écrivant (17.564) et en continuant la majoration en remplaçant  $f(h)$  par (17.562), nous obtenons

$$\begin{aligned} f(0) &\leq f(h) (1 + 2\|K + h\| \|A_1 h\| + q_1(h))^n \\ &\leq f(0) (1 + 2\|K\| \|Ah\| + q(h))^n (1 + 2\|K + h\| \|A_1 h\| + q_1(h))^n. \end{aligned} \quad (17.565)$$

Vu que nous avons dans l'idée de prendre des  $h$  petits (en norme) et que les parenthèses tendent vers 1 lorsque  $h$  est petit, nous pouvons supposer qu'elles sont non nulles et allègrement les passer au dénominateur. Nous divisons donc tout par le coefficient de  $f(h)$  :

$$\frac{f(0)}{(1 + 2\|K + h\| \|A_1 h\| + q_1(h))^n} \leq f(h) \leq f(0) (1 + 2\|K\| \|Ah\| + q(h))^n \quad (17.566)$$

Nous utilisons la règle de l'étau<sup>71</sup> pour conclure que

$$\lim_{h \rightarrow 0} \text{vol}(\mathcal{E}_h) = \text{vol}(\mathcal{E}_0). \quad (17.567)$$

- (4) Continuité Si  $a \in \mathbb{R}^n$  nous pouvons appliquer la limite (17.567) pour l'ellipsoïde  $\mathcal{E}_a$  au lieu de  $\mathcal{E}_0$ , c'est-à-dire en partant du compact  $K + a$  au lieu de  $K$ . Cela donne

$$\lim_{h \rightarrow 0} \text{vol}(\mathcal{E}_{a+h}) = \text{vol}(\mathcal{E}_a). \quad (17.568)$$

Donc  $f$  est continue sur  $\mathbb{R}^n$ .

### (5) Compact

Nous avons en hypothèse que  $K$  est d'intérieur non vide. Il existe donc  $a \in K$  et  $r > 0$  tel que  $B(a, r) \subset K$ . Soit l'hyperplan affine  $H$  passant par  $a$  et perpendiculaire à  $a$ . Nous notons  $S = H \cap B(a, r)$ . Nous avons évidemment  $S \subset K$  et donc  $S \subset \mathcal{E}_0$ . Vu que l'ellipsoïde  $\mathcal{E}_0$  est convexe et contient  $K$ , tout le cône de base  $S$  et de hauteur  $a$  est dans  $\mathcal{E}_0$ .

Sans rentrer dans les détails du calcul du volume de ce cône, son volume tend vers l'infini si  $a$  tend vers l'infini<sup>72</sup>

Nous avons donc une majoration

$$\text{vol}(\mathcal{E}_h) \geq \text{volume du cône de base } S \text{ et de hauteur } a + h. \quad (17.569)$$

Dès que ce volume est plus grand que  $\text{vol}(\mathcal{E}_0)$ , il est impossible que  $\text{vol}(\mathcal{E}_h) \leq \text{vol}(\mathcal{E}_0)$ . Soit  $R > 0$  tel que pour tout  $\|h\| \geq R$ , le volume du cône soit plus grand que  $\text{vol}(\mathcal{E}_0)$ . Le minimum de  $f$  est certainement dans  $B(0, R)$  parce que au moins

$$\text{vol}(\mathcal{E}_0) \in f(B(0, R)). \quad (17.570)$$

### (6) Minimum

Nous considérons la fonction

$$\begin{aligned} f: \overline{B(0, R)} &\rightarrow \mathbb{R}^+ \\ h &\mapsto \text{vol}(\mathcal{E}_h). \end{aligned} \quad (17.571)$$

Elle est continue sur un compact, et le théorème de Weierstrass 7.196 nous dit alors qu'elle atteint son minimum. Il n'y a cependant pas unicité de la valeurs de  $h$  pour laquelle  $f(h)$  est minimum.

71. Théorème 12.246.

72. Topologie du complété en un point, j'imagine que vous voyez de quoi il en retourne. Sinon c'est la définition 7.185, et il faut sans doute adapter le lemme 12.87 au cas de  $\mathbb{R}^n$  pour tout faire dans les règles.

(7) **Envoi**

Soit  $s \in \mathbb{R}^n$  tel que  $f(s)$  soit minimal. Cela signifie que parmi tous les ellipsoïdes centrés en zéro et contenant des ensembles de la forme  $K + h$ , le plus petit est celui qui contient  $K + s$ . Soit  $\mathcal{E}_s$  cet ellipsoïde.

Je prétend que  $\mathcal{E}_s - s$  est le plus petit ellipsoïde contenant  $K$ , tout centres confondus. En effet, si  $\mathcal{F}$  est un ellipsoïde centré en  $a$  et contenant  $K$ , alors  $\mathcal{F} - a$  est un ellipsoïde centré en zéro et contenant  $K - a$ . Nous avons alors

$$\text{vol}(\mathcal{F} - a) \geq \text{vol}(\mathcal{E}_s) = \text{vol}(\mathcal{E}_s - s). \quad (17.572)$$

L'invariance de la mesure par translation est le théorème 14.252.

□

## 17.16 Prolongement de fonctions

LEMooTUQIooEyTLBa

**Lemme 17.124** ([1]).

Soient un espace topologique  $X$  et un ouvert  $V \subset X$ . Nous considérons une fonction  $\phi \in C_c^0(V, \mathbb{C})$ . Alors l'extension

$$h: X \rightarrow \mathbb{C} \\ x \mapsto \begin{cases} \phi(x) & \text{si } x \in V \\ 0 & \text{sinon} \end{cases} \quad (17.573)$$

est continue sur  $X$ .

LemdCOMQM

**Lemme 17.125.**

Soit  $E$ , un espace vectoriel normé complet et  $(A_n)$  une suite emboîtée de fermés non vides dont le diamètre tend vers zéro. Alors l'intersection  $\bigcap_{n \in \mathbb{N}} A_n$  contient exactement un point.

*Démonstration.* Si l'intersection contenait deux points distincts  $a$  et  $b$ , alors nous aurions pour tout  $n$  la majoration  $\text{diam}(A_n) \geq \|a - b\|$  qui ne dépend pas de  $n$ . Cela contredirait la limite.

Soit une suite  $(x_n)$  avec  $x_k \in A_k$  pour tout  $k \in \mathbb{N}$ . C'est une suite de Cauchy. En effet si  $\epsilon > 0$ , considérons  $N$  tel que  $\text{diam}(A_N) < \epsilon$ . Dans ce cas dès que  $n, m > N$  nous avons  $x_n, x_m \in A_N$  et donc  $\|x_n - x_m\| \leq \epsilon$ . La suite  $x_n$  converge donc vers un élément dans  $E$ .

Nous devons montrer que  $x \in A_k$  pour tout  $k$ . La queue de suite  $(x_n)_{n \geq k}$  est une suite de Cauchy dans  $A_k$  qui converge donc vers un élément de  $A_k$  (ici nous utilisons le fait que  $A_k$  est fermé). Par unicité de la limite, cette dernière doit être  $x$ . Par conséquent  $x \in \bigcap_{n \in \mathbb{N}} A_n$ . □

ThoCaMpKO

**Théorème 17.126** ([591]).

Soient  $X$  et  $Y$  des espaces vectoriels normés. Pour une application linéaire  $f: X \rightarrow Y$ , les assertions suivantes sont équivalentes :

- (1)  $f$  est continue sur  $X$ ,
- (2)  $f$  est continue en un point de  $X$ ,
- (3)  $f$  est bornée.

PropTTiRgAq

**Proposition 17.127** ([592]).

Soit un espace normé  $X$ , un espace de Banach  $F$  et une partie dense  $A$  de  $X$ . Si l'application linéaire

$$f: (A, \|\cdot\|_X) \rightarrow F \quad (17.574)$$

est continue<sup>73</sup>, alors il existe une unique application linéaire continue  $\tilde{f}: X \rightarrow F$  prolongeant  $f$ . De plus  $\|\tilde{f}\| = \|f\|$ .

<sup>73</sup>. Nous avons bien mis sur  $A$  la topologie induite de  $X$ . Notons que ce n'est pas toujours celle qui est la plus naturelle sur  $A$ .



*Démonstration.* Soient  $x \in X$  et  $(A_n)_{n \geq 0}$  la suite d'ensembles définie par

$$A_n = \{y \in A \text{ tel que } \|x - y\| \leq 2^{-n}\}. \quad (17.575)$$

Étant donné que  $A$  est dense, ces ensembles sont tous non vides. De plus  $\text{diam } A_n \rightarrow 0$  parce que si  $y, y' \in A_n$  alors

$$\|y - y'\| \leq \|y - x\| + \|x - y'\| \leq 2^{-n+1}. \quad (17.576)$$

Vu que  $f$  est bornée, la suite d'ensembles  $f(A_n)$  est une suite emboîtée d'ensembles non vides de  $F$ . De plus leur diamètre tend vers zéro. En effet si  $z, z' \in f(A_n)$ , nous posons  $z = f(y)$ ,  $z' = f(y')$  et nous avons

$$\|z - z'\| \leq \|f(y) - f(x)\| + \|f(x) - f(y')\| \leq \|f\|(\|y - x\| + \|x - y'\|), \quad (17.577)$$

ce qui montre que  $\text{diam } f(A_n) \leq \|f\|2^{-n+1}$ . Notons que nous avons utilisé la linéarité de  $f$ . Par le lemme 17.125, l'intersection  $\bigcap_{n \in \mathbb{N}} \overline{f(A_n)}$  contient exactement un point. Nous posons

$$S(x) = \bigcap_{n \in \mathbb{N}} \overline{f(A_n)}. \quad (17.578)$$

Nous allons montrer que l'application  $x \mapsto S(x)$  ainsi définie est l'application que nous cherchons.

Nous commençons par montrer que pour toute suite  $y_k \rightarrow x$  avec  $y_k \in A$  nous avons

$$f(y_k) \rightarrow S(x). \quad (17.579)$$

Pour cela nous considérons  $n_0 \in \mathbb{N}$  et  $k_0$  tel que  $y_{k_0} \in A_{n_0}$ . Avec cela nous avons

$$\|f(y_k) - S(x)\| \leq \text{diam}(A_{n_0}) \leq \|f\|2^{-n_0+1}. \quad (17.580)$$

Pour montrer que  $S$  est linéaire, nous considérons deux suites dans  $A$  :  $y_k \rightarrow x$  et  $y'_k \rightarrow x'$  ainsi que la somme  $y_k + y'_k \rightarrow x + x'$ . Nous écrivons la relation (17.579) pour ces trois suites :

$$f(y_k) \rightarrow S(x) \quad (17.581a)$$

$$f(y'_k) \rightarrow S(x') \quad (17.581b)$$

$$f(y_k + y'_k) \rightarrow S(x + x'). \quad (17.581c)$$

Cependant, étant donné que  $f$  est linéaire, pour tout  $k$  nous avons  $f(y_k + y'_k) = f(y_k) + f(y'_k)$  et par conséquent

$$f(y_k + y'_k) \rightarrow S(x) + S(x'). \quad (17.582)$$

Par unicité de la limite,  $S(x + x') = S(x) + S(x')$ . Le même genre de raisonnement montre que  $S(\lambda x) = \lambda S(x)$ . L'application  $S$  est donc linéaire.

En ce qui concerne la continuité, nous avons

$$\|S(x)\| = \lim \|f(y_k)\| \leq \|f\| \lim \|y_k\| = \|f\|\|x\|, \quad (17.583)$$

donc  $\|S\| \leq \|f\|$ , c'est-à-dire que  $S$  est borné et donc continue parce que linéaire (théorème 17.126).

Nous montrons maintenant que  $S$  prolonge  $f$ . Si  $x \in A$ , alors nous avons  $\bigcap_{n \in \mathbb{N}} f(A_n) = f(x)$ , et donc  $S(x) = f(x)$ . Cela montre du même coup que  $\|f\| \leq \|S\|$  et que par conséquent  $\|f\| = \|S\|$ .

Passons à la partie sur l'unicité. Soient donc  $S$  et  $T$  deux prolongements continus de  $f$  sur  $X$ . Soient  $x \in X$  et  $x_n \rightarrow x$  une suite dans  $A$ . Par continuité nous avons  $T(x_n) \rightarrow T(x)$  et  $S(x_n) \rightarrow S(x)$ . Étant donné que par ailleurs pour tout  $n$  nous avons  $S(x_n) = T(x_n)$ , l'unicité de la limite montre que  $T(x) = S(x)$ .  $\square$



### 17.16.1 Encore du prolongement

Dans la même veine que la proposition 17.127 nous avons ce résultat.

ThoPVFQMj

**Théorème 17.128** ([593]).

Soient  $E$  et  $F$ , deux espaces métriques complets ainsi que  $A$  dense dans  $E$ . Si  $u: A \rightarrow F$  est uniformément continue, alors elle se prolonge de façon unique en une fonction continue  $\tilde{u}: E \rightarrow F$ . De plus ce prolongement est uniformément continu.

*Démonstration.* Soit  $x \in E \setminus A$  et une suite  $(x_n)$  contenue dans  $A$  et convergente vers  $x$ . Nous voulons définir

$$\tilde{u}(x) = \lim_{n \rightarrow \infty} u(x_n) \quad (17.584)$$

mais pour ce faire nous devons prouver que la suite  $(u(x_n))$  converge dans  $F$  et que la limite ne dépend pas de la suite choisie parmi les suites de  $A$  qui convergent (dans  $E$ ) vers  $x$ .

Commençons par montrer que  $(u(x_n))$  est de Cauchy dans  $F$ . Pour cela nous prenons  $\epsilon > 0$  et  $\eta > 0$  telle que  $d_E(a, b) < \eta$  implique  $d_F(u(a), u(b)) < \epsilon$  (uniforme continuité de  $u$ ). Après, il suffit de choisir  $N$  tel que pour tout  $n, m > N$  nous ayons  $d(x_m, x_n) < \eta$  (parce que  $x_n$  est de Cauchy). Avec tout ça nous avons

$$d_F(u(x_m), u(x_n)) < \epsilon, \quad (17.585)$$

ce qui signifie que  $(u(x_n))$  est de Cauchy et donc convergente dans  $F$ .

Nous voulons montrer maintenant que si  $(x_n)$  et  $(y_n)$  sont deux suites dans  $A$  convergentes vers  $x$  alors  $\lim_{n \rightarrow \infty} u(x_n) = \lim_{n \rightarrow \infty} u(y_n)$ . Pour cela nous considérons la suite  $z = (x_1, y_1, x_2, y_2, \dots)$ . Nous avons évidemment  $z_n \rightarrow x$ , et donc  $u(z_n)$  converge dans  $F$  par ce qui a été dit plus haut. Mais  $u(x_n)$  et  $u(y_n)$  en sont deux sous-suites convergentes. Donc leurs limites sont égales.

Il reste à montrer que ce  $\tilde{u}$  est continue et uniformément continue. Pour cela nous utilisons le module de continuité et le lemme 10.344. Étant donné que  $\tilde{u}$  prolonge  $u$  nous avons

$$\omega_{\tilde{u}}(h) \geq \omega_u(h). \quad (17.586) \quad \text{EqFRYgON}$$

Soient  $h > 0$  et  $\epsilon > 0$ ; soient aussi  $x, y \in E$  tels que  $d(x, y) < h$ . Nous prenons des suites  $(a_n) \rightarrow x$  et  $(b_n) \rightarrow y$  tout en choisissant  $n$  assez grand pour avoir  $d_E(a_n, b_n) < h$ . Nous avons

$$d_F(\tilde{u}(x), \tilde{u}(y)) \leq d_F(\tilde{u}(x), u(a_n)) + d(u(a_n), u(b_n)) + d_F(u(b_n), \tilde{u}(y)). \quad (17.587)$$

Si  $n$  est assez grand, par construction de  $\tilde{u}$ , le premier et le dernier terme sont plus petits que  $\epsilon$ . Par définition du module de continuité nous avons d'autre part  $d_F(u(a_n), u(b_n)) \leq \omega_u(h)$ . Du coup

$$d_F(\tilde{u}(x), \tilde{u}(y)) \leq \omega_u(h) + 2\epsilon. \quad (17.588)$$

Si nous prenons le supremum sur les  $x$  et  $y$  vérifiant  $d_E(x, y) < h$ , à gauche nous obtenons  $\omega_{\tilde{u}}(h)$  tandis que le membre de droite ne dépend pas de  $x$  et  $y$ . Donc pour tout  $\epsilon$ , nous avons

$$\omega_{\tilde{u}}(h) \leq \omega_u(h) + 2\epsilon. \quad (17.589)$$

En comparaison avec (17.586), nous trouvons

$$\omega_{\tilde{u}}(h) \leq \omega_u(h). \quad (17.590)$$

Les fonctions  $u$  et  $\tilde{u}$  ayant le même module de continuité, le lemme 10.344 nous enseigne que l'une est uniformément continue si et seulement si l'autre l'est. Vu que  $u$  est uniformément continue par hypothèse, le prolongement  $\tilde{u}$  est uniformément continu.  $\square$

DEFooMHQMooEvLccV

**Définition 17.129.**

Un **plongement** de l'espace topologique  $X$  dans  $Y$  est une application  $f: X \rightarrow Y$  telle que  $f: X \rightarrow f(X)$  soit un homéomorphisme.

**Théorème 17.130** (Extensiton des isométries).

Soit  $\tilde{M}$  un espace métrique complet et une application isométrique

$$f: A \rightarrow \tilde{M} \quad (17.591)$$

où  $A$  est une partie dense d'un espace métrique  $M$  (pas spécialement complet). Alors  $f$  accepte une unique extension isométrique

$$\tilde{f}: M \rightarrow \tilde{M} \quad (17.592)$$

Supposons de plus que  $M$  soit complet<sup>74</sup>. Alors  $\tilde{f}: M \rightarrow \tilde{M}$  est une bijection si et seulement si  $f(A)$  est dense dans  $\tilde{M}$ .

*Démonstration.* Nous commençons par prouver l'unicité. Soient  $\tilde{f}_1$  et  $\tilde{f}_2$  deux extensions de  $f$  et  $x \in M$ . Si  $(a_n)$  est une suite dans  $A$  convergeant vers  $x$  (possible parce que  $A$  est dense dans  $M$ ), alors nous avons

$$\tilde{f}_1(a_n) = \tilde{f}_2(a_n) \quad (17.593)$$

et donc  $\tilde{f}_1(x) = \tilde{f}_2(x)$  par continuité (une application isométrique est continue (proposition 7.306)).

Nous démontrons à présent l'existence.

- (1) **Construction de  $\tilde{f}$**  Soient  $x \in M$  et  $(a_n)$  une suite dans  $A$  qui converge vers  $x$ . Nous définissons

$$\tilde{f}(x) = \lim_{k \rightarrow \infty} f(a_k). \quad (17.594) \quad \text{EqHEamboy}$$

Note : nous pouvons prouver que cette définition ne dépend pas du choix de la suite  $(a_n)$  convergeant vers  $x$ , mais ce serait superflu parce que nous avons déjà prouvé l'unicité de  $\tilde{f}$ . Par contre nous devons expliquer pourquoi la limite du membre de droite de (17.594) existe dans  $\tilde{M}$ . D'abord la suite  $(a_n)$  est de Cauchy parce qu'elle est convergente (attention :  $M$  n'étant pas complet le fait d'être de Cauchy n'implique pas la convergence). Donc, étant donné que  $f$  est une isométrie, la suite  $(f(a_n))$  est de Cauchy dans  $\tilde{M}$ . Or ce dernier étant complet, la suite des images converge.

Montrons que cette application  $\tilde{f}: M \rightarrow \tilde{M}$  répond à la question.

- (2)  **$\tilde{f}$  est isométrique**

Soient  $a, b \in M$  et des suites dans  $A$  convergeant vers eux :  $a_n \rightarrow a$ ,  $b_n \rightarrow b$ . Nous avons, par continuité de l'application distance,

$$d(\tilde{f}(a), \tilde{f}(b)) = \lim_{k \rightarrow \infty} d(\tilde{f}(a_k), \tilde{f}(b)) \quad (17.595a)$$

$$= \lim_{k \rightarrow \infty} \lim_{l \rightarrow \infty} d(\tilde{f}(a_k), \tilde{f}(b_l)) \quad (17.595b)$$

$$= \lim_{k \rightarrow \infty} \lim_{l \rightarrow \infty} d(f(a_k), f(b_l)) \quad (17.595c)$$

$$= \lim_{k \rightarrow \infty} \lim_{l \rightarrow \infty} d(a_k, b_l) \quad (17.595d)$$

$$= d(a, b). \quad (17.595e)$$

Cela prouve que  $\tilde{f}$  est une isométrie.

Pour la suite nous supposons que  $M$  est complet. Notons tout de suite que  $\tilde{f}$  est injective parce qu'elle est isométrique.

- (3) **Bijection implique dense**

Nous supposons que  $\tilde{f}: M \rightarrow \tilde{M}$  est une bijection. Notons que  $f: A \rightarrow \tilde{M}$  a son image dans  $\tilde{M}$ . Donc  $f(A) \subset \tilde{M}$ .

Étant donné que  $\tilde{f}$  a pour image des limites de suites dans  $f(A)$ , l'image de  $\tilde{f}$  est contenue dans  $\overline{f(A)}$ , c'est-à-dire  $\tilde{f}(M) \subset \overline{f(A)}$ . Donc si  $\tilde{f}$  est surjective, c'est que  $\tilde{M} \subset \overline{f(A)}$ . Au final nous avons l'égalité  $\tilde{f}(A) = \tilde{M}$ . Cela prouve que si  $\tilde{f}$  est bijective, alors  $f(A)$  est dense dans  $\tilde{M}$ .

74. Il me semble que cette hypothèse manque dans [594].

(4) **Dense implique bijective**

Nous supposons que  $\overline{f(A)} = \tilde{M}$  et nous devons prouver que  $\tilde{f}$  est surjective. Soient  $x \in \tilde{M}$  et  $f(a_n)$  une suite dans  $f(A)$  qui converge vers  $x$ ; une telle suite existe parce que  $f(A)$  est dense dans  $\tilde{M}$ . Cette suite est de Cauchy dans  $\tilde{M}$  parce que dans un espace métrique, une suite convergente est de Cauchy. La suite  $(a_n)$  est elle-même également de Cauchy parce que

$$d(a_n, a_m) = d(f(a_n), f(a_m)). \quad (17.596)$$

Étant donné que  $(a_n)$  est de Cauchy dans  $M$ , elle converge vers un élément que nous nommons  $a \in M$ . Par continuité de  $f$  nous avons alors

$$\tilde{f}(a) = \lim_{k \rightarrow \infty} f(a_k) = x. \quad (17.597)$$

Cela prouve que  $x$  est bien dans l'image de  $\tilde{f}$  et donc que  $\tilde{f}$  est surjective. □

## 17.17 Complétion d'un espace métrique

**Lemme 17.131** ([595]).

Soient des espaces métriques  $(X, d)$  et  $(Y, \rho)$ . Soit une application isométrique  $f: X \rightarrow Y$ .

- (1) Si  $(x_n)$  est de Cauchy dans  $X$ , alors  $(f(x_n))$  est de Cauchy dans  $Y$ .
- (2) Si  $(a_n)$  et  $(b_n)$  sont deux suite des Cauchy dans  $X$  ayant la même limite, alors  $(f(a_n))$  et  $(f(b_n))$  ont même limite.

*Démonstration.* Soit  $\epsilon > 0$ . Soit  $N_\epsilon$  tel que si  $p, q \geq N_\epsilon$ , nous avons  $d(x_p, x_q) \leq \epsilon$ . Pour ce même  $N_\epsilon$ , vu que  $f$  est isométrique,

$$\rho(f(x_p), f(x_q)) = d(x_p, x_q) \leq \epsilon. \quad (17.598)$$

Cela prouve (1). Nous passons à (2).

Supposons  $a_n \xrightarrow{X} x$  et  $b_n \xrightarrow{X} x$ , et nous considérons  $y_a$  et  $y_b$  les limites  $f(a_n) \xrightarrow{Y} y_a$  et  $f(b_n) \xrightarrow{Y} y_b$ . Nous avons

$$\rho(y_a, y_b) \leq \rho(y_a, f(a_n)) + \underbrace{\rho(f(a_n), f(b_n))}_{=d(a_n, b_n)} + \rho(f(b_n), y_b) \quad (17.599a)$$

$$\leq \rho(y_a, f(a_n)) + d(a_n, x) + d(x, b_n) + \rho(f(b_n), y_b). \quad (17.599b)$$

À droite, les quatre termes tendent vers zéro lorsque  $n \rightarrow \infty$ . Donc  $\rho(y_a, y_b) = 0$ , et donc  $y_a = y_b$ . □

**Lemme 17.132** ([595]).

Soient des espaces métriques  $(X, d)$  et  $(Y, \rho)$  ainsi que  $D$  dense dans  $X$ . Nous supposons que  $Y$  est dense et nous considérons une application isométrique  $f: D \rightarrow Y$ .

Alors il existe une unique isométrie  $F: X \rightarrow Y$  qui étend  $f$ .

*Démonstration.* En plusieurs points.

- (1) **Construction de  $F$**  Nous commençons par l'existence. Soit  $x \in X$ . Étant donné que  $D$  est dense dans  $X$ , il existe une suite  $(x_n)$  dans  $D$  telle que  $x_n \xrightarrow{X} x$ . Nous déduisons du lemme 17.131 que  $(f(x_n))$  est de Cauchy dans  $Y$ , et même que la limite ne dépend pas de la suite  $(x_n)$  convergeant vers  $x$ . Bref, nous pouvons définir

$$\begin{aligned} F: X &\rightarrow Y \\ x &\mapsto \lim_{n \rightarrow \infty} f(x_n) \end{aligned} \quad (17.600)$$

où  $(x_n)$  est n'importe quelle suite dans  $D$  qui converge vers  $x$ . Si  $x \in D$ , alors il suffit de prendre la suite constante  $x_n = x$  pour obtenir que  $F(x) = f(x)$ .

(2) **F est une isométrie**

Soient  $x, y \in X$  ainsi que deux suites de Cauchy  $x_n \xrightarrow{X} x$  et  $y_n \xrightarrow{X} y$ . Soit  $\epsilon > 0$ . Si  $n$  est assez grand,

$$d(x, y) \leq d(x, x_n) + d(x_n, y_n) + d(y, y_n) \quad (17.601a)$$

$$\leq 2\epsilon + \rho(f(x_n), f(y_n)) \quad f \text{ est isom.} \quad (17.601b)$$

$$\leq 2\epsilon + \rho(f(x_n), F(x)) + \rho(F(x), F(y)) + \rho(F(y), f(y_n)) \quad (17.601c)$$

$$\leq 3\epsilon + \rho(F(x), F(y)). \quad (17.601d)$$

Donc nous avons l'inégalité

$$d(x, y) \leq \rho(F(x), F(y)). \quad (17.602)$$

En partant maintenant de  $\rho(F(x), F(y))$  et en effectuant les mêmes majorations, nous trouvons l'inégalité dans l'autre sens :  $\rho(F(x), F(y)) \leq d(x, y)$ , de telle sorte que

$$d(x, y) = \rho(F(x), F(y)), \quad (17.603)$$

ce qui signifie que  $F$  est une isométrie.

(3) **Unicité** Soient deux isométries  $F: X \rightarrow Y$  et  $F': X \rightarrow Y$  telles que  $F = F'$  sur  $D$ . Nous devons prouver que  $F(x) = F'(x)$  pour tout  $x \in X$ .

Soient donc  $x \in X$  ainsi qu'une suite  $(x_n)$  dans  $D$  telle que  $x_n \xrightarrow{X} x$ . Nous avons alors

$$\rho(F(x), F'(x)) \leq \rho(F(x), F(x_n)) + \underbrace{\rho(F(x_n), F'(x_n))}_{=0} + \rho(F'(x_n), F'(x)) \quad (17.604a)$$

$$= d(x, x_n) + d(x_n, x). \quad (17.604b)$$

Donc pour tout  $n$  nous avons  $\rho(F(x), F'(x)) \leq d(x, x_n) + d(x_n, x)$ . En faisant  $n \rightarrow \infty$  nous trouvons  $\rho(F(x), F'(x)) \leq 0$  et donc  $\rho(F(x), F'(x)) = 0$ .

□

DEFooXHQFooNBIAvy

**Définition 17.133** ([595]).

Soit un espace métrique  $(X, d)$ . Une **complétion** de  $(X, d)$  est un triple  $(\hat{X}, \hat{d}, j)$  où

- (1) L'espace  $(\hat{X}, \hat{d})$  est un espace métrique complet,
- (2) L'application  $j: X \rightarrow \hat{X}$  est une isométrie,
- (3) La partie  $j(X)$  est dense dans  $\hat{X}$ .

THOooRLYBooCVBjoP

**Théorème 17.134** (Unicité de la complétion[595]).

Soient  $(X, d)$  un espace métrique ainsi que deux complétions<sup>75</sup>  $(X_1, d_1, j_1)$ ,  $(X_2, d_2, j_2)$ . Alors les espaces  $(X_1, d_1)$  et  $(X_2, d_2)$  sont isométriques.

*Démonstration.* Nous considérons l'application  $f = j_2 \circ j_1^{-1}: j_1(X) \rightarrow X_2$ . Cette application est isométrique (parce que  $j_1$  et  $j_2$  le sont), l'espace  $X_2$  est complet et  $j_1(X)$  est dense dans  $X_1$ . Le lemme 17.132 nous permet de considérer une isométrie  $F: X_1 \rightarrow X_2$ . Nous devons prouver que  $F$  est bijective.

L'application  $F$  est injective parce qu'elle est isométrique. En ce qui concerne la surjectivité, soit  $y \in X_2$ . Étant donné que  $j_2(X)$  est dense dans  $X_2$ , il existe une suite  $(y_n)$  dans  $X_2$  telle que  $y_n \xrightarrow{X_2} y$ . Pour chaque  $n$  nous choisissons  $x_n \in X$  tel que  $j_2(x_n) = y_n$ . Nous posons  $z_n = j_1(x_n) \in X_1$ .

Vu que  $(y_n)$  est de Cauchy et que  $j_1$  est isométrique, la suite  $(z_n)$  est de Cauchy. Étant donné que  $X_1$  est complet, la suite  $(z_n)$  a une limite dans  $X_1$ . Nous pouvons donc poser  $z = \lim_{n \rightarrow \infty} z_n \in X_1$ .

Nous prouvons à présent que  $F(z) = y$ . L'application  $F: X_1 \rightarrow X_2$  est isométrique et donc continue ; elle commute donc avec la limite et

$$\lim_{n \rightarrow \infty} d_2(F(z_n), y) = d_2(F(z), y). \quad (17.605)$$

---

75. Définition 17.133.

Mais  $j_1(x) \in j_1(X)$ , et nous savons que  $F = f$  sur  $j_1(X)$ . Donc

$$F(j_1(x_n)) = f(j_1(x_n)) = j_2(x_n). \quad (17.606)$$

Donc nous avons le calcul

$$d_2(F(z), y) = \lim_{n \rightarrow \infty} d_2(F(z_n), y) \quad (17.607a)$$

$$= \lim_{n \rightarrow \infty} d_2(F(j_1(x_n)), y) \quad (17.607b)$$

$$= \lim_{n \rightarrow \infty} d_2(j_2(x_n), y) \quad (17.607c)$$

$$= \lim_{n \rightarrow \infty} d_2(y_n, y) \quad (17.607d)$$

$$= 0. \quad (17.607e)$$

Cela prouve que  $F(z) = y$  et donc que  $F$  est surjective.  $\square$

Une conséquence du théorème de prolongement est le théorème suivant qui permet de compléter un espace métrique.

ThoKHTQJXZ

**Théorème 17.135** (Complétion d'un espace métrique[596, 594, 595]).

Plus précisément, soit  $(M, d)$  un espace métrique. Il existe un espace métrique complet  $\tilde{M}$  muni d'un plongement<sup>76</sup> isométrique  $\varphi: M \rightarrow \tilde{M}$  tel que  $\varphi(M)$  soit dense dans  $\tilde{M}$ .

Ce complété de  $M$  est unique au sens suivant. Si  $\tilde{M}_1$  et  $\tilde{M}_2$  sont deux espaces métriques complets munis de plongements isométriques  $f_i: M \rightarrow \tilde{M}_i$  dont les images sont denses, alors il existe une bijection isométrique  $\phi: \tilde{M}_1 \rightarrow \tilde{M}_2$  telle que  $\phi \circ f_1 = f_2$ .

*Démonstration.* L'unicité est déjà le théorème 17.134. Nous faisons l'existence.

(1) **Existence**

Soit  $C_M$  l'ensemble des suites de Cauchy de  $M$ . Nous définissons

$$\begin{aligned} f: C_M \times C_M &\rightarrow \mathbb{R} \\ u, v &\mapsto \lim_{n \rightarrow \infty} d(u_n, v_n). \end{aligned} \quad (17.608)$$

Notre première tâche est de nous assurer que cela est bien défini, c'est-à-dire que la limite existe toujours. En effet, si  $u$  et  $v$  sont des suites de Cauchy dans  $M$ , nous avons

$$|d(u_n, v_n) - d(u_m, v_m)| \leq |d(u_n, v_n) + d(u_n, v_m)| + |d(u_n, v_m) - d(u_m, v_m)| \quad (17.609a)$$

$$\leq d(v_n, v_m) + d(u_n, u_m) \quad \text{lem. 7.141} \quad (17.609b)$$

$$\leq 2\epsilon \quad (17.609c)$$

dès que  $m$  et  $n$  sont assez grand. Cela prouve que la suite  $n \mapsto d(u_n, v_n)$  est de Cauchy dans  $\mathbb{R}$ . Par complétude de  $\mathbb{R}$ , elle converge<sup>77</sup>.

Nous considérons la relation d'équivalence  $u \sim v$  si et seulement si  $f(u, v) = 0$ . Nous posons  $\tilde{M} = C_M / \sim$  et nous y mettons la distance

$$d([u], [v]) = f(u, v) \quad \text{EqDDI/NBNE} \quad (17.610)$$

et nous devons encore vérifier que cela est bien défini. Prenons  $u' \sim u$  et  $v' \sim v$ . Alors nous avons

$$d(u'_n, v'_n) \leq d(u'_n, u_n) + d(u_n, v_n) + d(v_n, v'_n), \quad (17.611)$$

76. Définition 17.129.

77. Ici nous utilisons la complétude de  $\mathbb{R}$ . Cette dernière doit donc être démontrée indépendamment, ce qui est fait dans le théorème 7.339. De plus nous ne pouvons pas définir  $\mathbb{R}$  comme étant le complété de  $\mathbb{Q}$  en utilisant ce théorème.

et donc

$$d(u', v') = \lim_{n \rightarrow \infty} d(u'_n, v'_n) \leq \lim_{n \rightarrow \infty} d(u_n, v_n) = d(u, v). \quad (17.612)$$

Le même argument en inversant les primes et les non primes montre l'inégalité inverse. Donc  $d(u, v) = d(u', v')$  dans  $C_M$ , et donc la distance (17.610) est bien définie sur  $\tilde{M}$ .

Afin de s'assurer que  $\tilde{M}$  répond bien à la question du théorème, il faut encore démontrer les points suivants :

- $M$  se plonge isométriquement dans  $\tilde{M}$ .
- l'image de  $M$  par le plongement est dense dans  $\tilde{M}$ .
- $\tilde{M}$  est complet.

Nous allons maintenant considérer l'application

$$\begin{aligned} \varphi: M &\rightarrow \tilde{M} \\ x &\mapsto \text{la classe de la suite constante } x. \end{aligned} \quad (17.613)$$

(1a) **Plongement isométrique** Nous allons montrer que cela est une isométrie bijective et que  $\varphi(M)$  est dense dans  $\tilde{M}$ . Le fait que  $\varphi$  soit bijective entre  $M$  et  $\varphi(M)$  est évident. C'est une isométrie parce que

$$d(\varphi(x), \varphi(y)) = \lim_{n \rightarrow \infty} d(\varphi(x)_n, \varphi(y)_n) = d(x, y). \quad (17.614)$$

(1b) **Densité**

Soit  $[u] \in \tilde{M}$ . Tous les termes  $u_n$  sont des éléments de  $M$ . Nous considérons la suite dans  $\varphi(M)$  donnée par

$$a_n = \varphi(u_n) \quad (17.615)$$

Chaque  $a_n$  est un élément <sup>78</sup> de  $\tilde{M}$ . Montrons que  $(a_n)$  converge dans  $\tilde{M}$  vers  $u$ . Nous avons

$$d(a_n, u) = \lim_{k \rightarrow \infty} d((a_n)_k, u_k) \quad (17.616a)$$

$$= \lim_{k \rightarrow \infty} d(u_n, u_k) \quad (17.616b)$$

$$= d(u_n, \ell) \quad (17.616c)$$

en notant  $\ell$  la limite de la suite  $(u_n)$ . Ici nous avons utilisé le fait que la fonction distance était continue pour l'inverser avec la limite, par le théorème 12.244. Nous avons alors

$$\lim_{n \rightarrow \infty} d(a_n, [u]) = \lim_{n \rightarrow \infty} d(u_n, \ell) = 0. \quad (17.617)$$

(1c) **Complétude** Nous passons maintenant à la preuve du fait que  $\tilde{M}$  est complet. Soit  $(y_n)$  une suite de Cauchy dans  $\tilde{M}$ . Soit  $\epsilon > 0$  ; nous définissons  $K(n)$  par

$$d((y_n)_k, (y_n)_l) < \epsilon \quad (17.618)$$

dès que  $k, l \geq K(n)$ . Cette définition fonctionne parce que pour chaque  $n$ ,  $y_n$  est une suite de Cauchy dans  $M$ . Nous posons

$$x_n = (y_n)_{K(n)} \in M \quad (17.619)$$

et nous allons montrer que  $(x_n)$  est de Cauchy dans  $M$  –donc est un élément de  $\tilde{M}$ – et que  $y_k \rightarrow (x_n)$  dans  $\tilde{M}$ .

Nous commençons par montrer que  $(x_n)$  est de Cauchy dans  $M$ . Nous avons

$$d(x_n, x_m) = d((y_n)_{K(n)}, (y_m)_{K(m)}) \quad (17.620a)$$

$$\leq d((y_n)_{K(n)}, (y_n)_l) + d((y_n)_l, (y_m)_l) + d((y_m)_l, (y_m)_{K(m)}) \quad (17.620b)$$

---

78. À partir de maintenant nous n'écrivons plus explicitement la classe d'équivalence.

Nous choisissons  $n, m$  tels que  $d(y_n, y_m) < \epsilon$ , ce qui nous permet de choisir  $l$  de telle façon à avoir  $d((y_n)_k, (y_m)_k) < \epsilon$  pour tout  $k \geq l$ . De plus, quitte à encore augmenter  $l$ , nous supposons que  $l > K(m)$  et  $l > K(n)$ . Avec ces choix nous voyons que  $d(x_n, x_m) < 3\epsilon$ , ce qui signifie que la suite  $(x_n)$  est de Cauchy dans  $M$ .

En ce qui concerne la convergence  $y_n \rightarrow (x)$ , on a

$$d(y_n, (x)) = \lim_{k \rightarrow \infty} d((y_n)_k, (y_k)_{K(k)}) \quad (17.621a)$$

$$\leq \lim_{k \rightarrow \infty} d((y_n)_k, (y_n)_l) + \lim_{k \rightarrow \infty} d((y_n)_l, (y_k)_l) + \lim_{k \rightarrow \infty} d((y_k)_l, (y_k)_{K(k)}) \quad (17.621b)$$

Nous devons trouver un  $n$  tel que si  $k$  est suffisamment grand, le tout est majoré par  $\epsilon$ . Voici nos choix :

- $n$  tel que  $d(y_n, y_m) < \epsilon$  dès que  $m \geq n$ ,
- $k > n$ ,
- $k > K(n)$ ,
- $l > k$ ,
- $l > K(k)$ ,
- $l$  suffisamment grand pour que  $d((y_n)_l, (y_k)_l) < \epsilon$ .

Avec tous ces choix, les trois termes de (17.621b) sont plus petits que  $\epsilon$ .

Ceci prouve que  $\tilde{M}$  est complet.

Unicité Cette partie de la preuve est encore à écrire. Si vous savez comment faire, écrivez-moi.

□

### 17.17.1 Principe des zéros isolés

**Théorème 17.136** (Principe des zéros isolés [597]).

Soit un ouvert  $\Omega \subset \mathbb{C}$ . Soient une fonction analytique<sup>79</sup>  $f: \Omega \rightarrow \mathbb{C}$  et  $a \in \Omega$ , un zéro non isolé de  $f$ . Alors  $f$  est nulle sur un voisinage de  $a$ .

*Démonstration.* Nous écrivons  $f$  sous la forme d'une série entière<sup>80</sup> autour de  $a$  :

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n \quad (17.622)$$

valable sur une boule  $B(a, r)$ . Soit  $c_m$  le premier coefficient non nul (si il n'existe pas c'est que  $f$  est nulle sur tout  $B(a, r)$  et alors le théorème est prouvé). Nous avons alors

$$f(z) = c_m (z-a)^m \left(1 + \sum_{k=1}^{\infty} d_k (z-a)^k\right) \quad (17.623)$$

avec  $d_k = c_{k+m}/c_m$ . Le rayon de convergence de la série  $\sum_k d_k (z-a)^k$  est le même que celui de (17.622) parce que la suite  $\frac{d_k r^{m+k}}{c_m}$  reste bornée (critère d'Abel, lemme 15.18). Si nous posons

$$g(z) = 1 + \sum_{k=1}^{\infty} d_k (z-a)^k, \quad (17.624)$$

alors  $g$  est une fonction continue et  $g(a) = 1$ . De plus

$$f(z) = c_m (z-a)^m g(z). \quad (17.625)$$

<sup>79</sup>. Définition 12.491.

<sup>80</sup>. Définition de fonction analytique 12.491.



Soit une suite  $(z_n)$  de zéros de  $f$  qui converge vers  $a$ . Étant donné que  $g$  est continue, nous devrions avoir  $\lim_{k \rightarrow \infty} g(z_k) = g(a) = 1$ , mais si  $f(z_k) = 0$  avec  $z_k \neq a$ , alors  $g(z_k) = 0$ . Cela est un paradoxe qui nous permet de conclure que si la suite  $z_n$  existe bien, alors  $f$  est identiquement nulle sur un voisinage, c'est-à-dire que tous les  $c_n$  sont nuls.  $\square$

CORooFBXXooZyfUQi

**Corolaire 17.137.**

Soit  $f$  une fonction analytique sur un ouvert connexe  $\Omega \subset \mathbb{C}$ . Si  $f$  s'annule sur un ouvert (non vide) de  $\Omega$ , alors  $f$  s'annule sur tout  $\Omega$ .

*Démonstration.* soit

$$N = \{z \in \Omega \text{ tel que } f = 0 \text{ sur un ouvert autour de } z\}. \quad (17.626)$$

Le fait que  $N$  soit ouvert est évident à partir de sa définition. Nous allons montrer que  $N$  est également fermé dans  $\Omega$ , et donc conclure que  $N = \Omega$ . Soit  $(z_n)$  une suite dans  $N$  convergente vers  $z \in \Omega$ . Étant donné que  $f(z_n) = 0$  et que  $f$  est continue, nous avons

$$f(z) = \lim_{n \rightarrow \infty} f(z_n) = 0, \quad (17.627)$$

ce qui fait de  $z$  un zéro non isolé de  $f$ . Par conséquent le principe des zéros isolés (théorème 17.136) nous enseigne que  $f$  s'annule dans un voisinage autour de  $z$ , c'est-à-dire que  $z \in N$ . L'ensemble  $N$  est donc fermé.  $\square$

ThoAVBCewB

**Théorème 17.138** (Principe du prolongement analytique[1]).

Soit  $\Omega$  un ouvert connexe de  $\mathbb{C}$ . Si deux fonctions analytiques coïncident sur un sous-ensemble  $D$  de  $\Omega$  contenant un point d'accumulation dans  $\Omega$ , alors elles sont égales sur  $\Omega$ .

*Démonstration.* Soient deux fonctions analytique  $f, g: \Omega \rightarrow \mathbb{C}$  égales sur  $D$ . Nous considérons la fonction  $h = f - g$ . Il existe dans  $D$  une suite  $(z_n)$  qui converge vers  $a \in \Omega$  et telle que  $f(z_n) = g(z_n)$ , c'est-à-dire  $h(z_n) = 0$  pour tout  $n$ .

Vu que  $h$  est continue nous avons aussi  $h(a) = 0$ . Le point  $a$  est donc un zéro non isolé de  $h$ . Le théorème 17.136 conclut que  $h = 0$  sur  $\Omega$ .  $\square$

## 17.18 Un petit extra

Ceci provient de notes de TP de l'université libre de Bruxelles autour des années 2000-2001[598].

LEMooXRMAooRADhOM

**Lemme 17.139** ([1]).

Soit une fonction  $f: \mathbb{R} \rightarrow \mathbb{R}$  telle que

$$(1) \quad f(1) = 1,$$

ItemExtrai

$$(2) \quad f(x + y) = f(x) + f(y) \text{ pour tout réels } x \text{ et } y.$$

ItemExtraii

Alors  $f(q) = q$  pour tout  $q \in \mathbb{Q}$ .

*Démonstration.* Vu que  $f(x) + f(y) = f(x + y)$ , nous avons, pour  $p \in \mathbb{N}$  :

$$\sum_{i=1}^p f\left(\frac{1}{p}\right) = f\left(\sum_{i=1}^p \frac{1}{p}\right) = f(1) = 1. \quad (17.628)$$

Donc  $f(1/p) = 1/p$ . Nous en déduisons tout de suite que  $p(p/q) = pf(1/q) = p/q$ .  $\square$

Si nous ajoutons l'hypothèse de continuité, nous avons un résultat plus fort.

LEMooYKCUooUiIvPJ

**Lemme 17.140** ([1]).

Soit une fonction  $f: \mathbb{R} \rightarrow \mathbb{R}$  telle que

$$(1) \quad f(1) = 1,$$



(2)  $f(x + y) = f(x) + f(y)$  pour tout réels  $x$  et  $y$ .

(3)  $f$  est continue.

Alors  $f$  est la fonction identité :  $f(x) = x$  pour tout  $x \in \mathbb{R}$ .

*Démonstration.* Nous savons déjà du lemme 17.139 que  $f(q) = q$  pour tout  $q \in \mathbb{Q}$ . Soient  $x \in \mathbb{R}$ , ainsi qu'une suite  $q_k$  dans  $\mathbb{Q}$  telle que  $q_k \xrightarrow{\mathbb{R}} x$ . Par continuité, nous avons

$$f(x) = \lim_{k \rightarrow \infty} f(q_k) = \lim_{k \rightarrow \infty} f(q_k) = \lim_{k \rightarrow \infty} q_k = x \quad (17.629)$$

parce que  $f(q_k) = q_k$  pour tout  $k$ .  $\square$

Une question naturelle qu'on peut alors se poser est la suivante :

Est-il possible de définir une fonction réelle non continue ayant les propriétés (1) et (2) ?

Nous savons qu'une application linéaire sur un espace vectoriel de dimension finie est continue<sup>81</sup>, mais qu'en dimension infinie, une application linéaire non continue est possible<sup>82</sup>.

Donc une manière de trouver une réponse positive à la question posée plus haut, serait de voir  $\mathbb{R}^n$  comme espace vectoriel de dimension infinie.

Nous nous souvenons que, grâce au lemme de Zorn<sup>83</sup>, tout espace vectoriel admet une base<sup>84</sup>. En particulier, l'ensemble des réels vu comme espace vectoriel sur  $\mathbb{Q}$  admet une base, i.e.  $\exists (e_i)_{i \in I}$  des éléments de  $\mathbb{R}$  tels que tout réel s'écrit comme combinaison linéaire à coefficients rationnels de ces  $e_i$ , i.e.

$$\forall r \in \mathbb{R}, \exists (\lambda_i)_{i \in I} \text{ des éléments de } \mathbb{Q} \text{ tels que } r = \sum_{i \in I} \lambda_i e_i. \quad (17.630)$$

C'est avec ça en main que nous allons faire le travail.

### Proposition 17.141.

Soient  $e_1 = 1$  ainsi qu'un irrationnel  $e_2$ . Soit une  $\mathbb{Q}$ -base complétée  $(e_i)_{i \in I}$  avec<sup>85</sup> de  $\mathbb{R}$ .

Nous définissons  $h: \mathbb{R} \rightarrow \mathbb{R}$  par

$$h(e_1) = e_1 \quad (17.631a)$$

$$h(e_2) = e_3 \quad (17.631b)$$

$$h(e_3) = e_2 \quad (17.631c)$$

$$h(e_i) = e_i \quad \forall i \in I \setminus \{1, 2, 3\}, \quad (17.631d)$$

et par  $\mathbb{Q}$ -linéarité :

$$h\left(\sum \lambda_i e_i\right) = \sum \lambda_i \alpha_i \quad (17.632)$$

pour tout  $(\lambda_i \in \mathbb{Q})_{i \in I}$ .

La fonction  $h$  vérifie

(1)  $h(1) = 1$

(2)  $h$  est  $\mathbb{Q}$ -linéaire, c'est-à-dire que

(2a)  $h(x + y) = h(x) + h(y)$  pour tout  $x, y \in \mathbb{R}$

(2b)  $h(\lambda x) = \lambda h(x)$  pour tout  $\lambda \in \mathbb{Q}$ .

(3)  $h$  n'est pas continue (pour la topologie usuelle sur  $\mathbb{R}$ ).

*Démonstration.* En plusieurs points

(1) **Pour (1)** C'est la condition  $h(e_1) = e_1$  en nous souvenant que  $e_1 = 1$ .

(2) **Pour (2a)** Nous décomposons  $x$  et  $y$  dans la base  $\{e_i\}_{i \in I}$ . Il existe  $J_x$  et  $J_y$  finis dans  $I$  tels que<sup>86</sup>  $x = \sum_{i \in J_x} x_i e_i$ ,  $y = \sum_{i \in J_y} y_i e_i$ . Vu que toutes les sommes sont finies, nous n'avons

81. Proposition 7.234.

82. Exemples 10.82 et 10.83.

83. Lemme 1.37.

84. Proposition 4.25.

85. Théorème de la base incomplète 4.26. Attention : nous ne prétendons pas que  $I$  est dénombrable. Par contre, par définition d'une base, les sommes sont toujours finies. Tout réel peut s'écrire comme une combinaison linéaire finie des  $e_i$  avec des coefficients rationnels.

86. Les vrais chasseurs de précisions diront que  $x_i$  est défini pour  $i \in I$ , mais uniquement non nul sur  $J_x$ .

ITEMooOXXMooJFSuuN

ITEMooAVNMooEsodLR

ITEMooDEJMooXtvrqi

ITEMooMHNZooBHQywD

aucun problème à écrire

$$x + y = \sum_{i \in J_x \cup J_y} (x_i + y_i) e_i \quad (17.633)$$

et donc

$$h(x + y) = \sum_i h(x_i + y_i) = \sum_i h(x_i) + \sum_i h(y_i) = h(x) + h(y). \quad (17.634)$$

- (3) Pour (2b) Même type de vérifications que pour (2a).  
 (4) Pour (3) Si  $h$  était continue, elle serait l'identité par le lemme 17.140. Or elle n'est pas l'identité parce que  $h(e_2) = e_3$ .

□

Donc nous avons trouvé une application  $\mathbb{Q}$ -linéaire  $h: \mathbb{R} \rightarrow \mathbb{R}$  qui n'est pas continue.



# Chapitre 18

## Trigonométrie, isométries

### 18.1 Trigonométrie

#### 18.1.1 Définitions, périodicité et quelques valeurs remarquables

**Proposition-Définition 18.1** (Définition du cosinus et du sinus).

PROPOoZXPVooBjONka

La série

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \quad (18.1)$$

définit une fonction  $\cos: \mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^\infty$ . Nous l'appelons **cosinus**.

La série

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad (18.2)$$

EQooCMRFooCTtpge

définit une fonction  $\sin: \mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^\infty$ . Nous l'appelons **sinus**.

*Démonstration.* La série entière définissant  $\cos(x)$  a pour coefficients

$$a_n = \begin{cases} 0 & \text{si } n \text{ est impair} \\ \frac{(-1)^{n/2}}{n!} & \text{si } n \text{ est pair.} \end{cases} \quad (18.3)$$

Nous pouvons la majorer par la série entière donnée par les coefficients

$$b_n = \begin{cases} 1/n! & \text{si } n \text{ est impair} \\ \frac{(-1)^{n/2}}{n!} & \text{si } n \text{ est pair.} \end{cases} \quad (18.4)$$

Quelle que soit la parité de  $k$  nous avons toujours

$$\left| \frac{b_{k+1}}{b_k} \right| = \frac{1}{k+1}, \quad (18.5)$$

de telle sorte que la formule d'Hadamard (15.98) nous donne  $R = \infty$  pour la série  $\sum_{k=0}^{\infty} b_k x^k$ . A fortiori<sup>1</sup> le rayon de convergence pour la série du cosinus est infini.

L'assertion concernant le sinus se démontre de même.

En ce qui concerne le fait que les fonctions sin et cos sont de classe  $C^\infty$  sur  $\mathbb{R}$ , il faut invoquer le corolaire 15.52.  $\square$

LEMooZMJTooJPnyfv

**Lemme 18.2** ([1]).

Nous avons  $\text{SUBEQooTTNNooXzApSM}$

$$\cos(0) = 1 \quad (18.6a)$$

$$\sin(0) = 0. \quad (18.6b)$$

---

1. Remarque 15.16.

ainsi que  $\left[ \begin{smallmatrix} \text{SUBEQSooFOGNooQrBxYc} \end{smallmatrix} \right]$

$$\cos(-x) = \cos(x) \quad (18.6\text{ca})$$

$$\sin(-x) = -\sin(x) \quad (18.6\text{cb})$$

$$(18.6\text{cc})$$

*Démonstration.* Par substitution directe dans les séries.  $\square$

$\left[ \begin{smallmatrix} \text{LEMooBBCAooHLWmno} \end{smallmatrix} \right]$

### Lemme 18.3.

*En ce qui concerne la dérivation, nous avons*

$$\sin' = \cos \quad (18.4\text{a})$$

$$\cos' = -\sin. \quad (18.4\text{b})$$

*Démonstration.* Il s'agit de se permettre de dériver terme à terme (proposition 15.50) les séries qui définissent le sinus et le cosinus.  $\square$

$\left[ \begin{smallmatrix} \text{LEMooAEFPooGSgOkF} \end{smallmatrix} \right]$

### Lemme 18.4.

*Les fonctions sinus et cosinus vérifient*

$$\cos^2(x) + \sin^2(x) = 1 \quad (18.5)$$

$\left[ \begin{smallmatrix} \text{EQooNYCZooApyyBd} \end{smallmatrix} \right]$

pour tout  $x \in \mathbb{R}$ .

*Démonstration.* Posons  $f(x) = \sin^2(x) + \cos^2(x)$  et dérivons :

$$f'(x) = 2 \sin(x) \cos(x) + 2 \cos(x)(-) \sin(x) = 0. \quad (18.6)$$

La fonction  $f$  est donc constante par le corolaire 12.221. Nous avons donc pour tout  $x$  :

$$f(x) = f(0) = \sin^2(0) + \cos^2(0) = 1. \quad (18.7)$$

Le dernier calcul s'obtient en substituant directement  $x$  par zéro dans les séries :  $\sin(0) = 0$  et  $\cos(0) = 1$ .  $\square$

## 18.1.2 Fonction puissance (pour les complexes)

La fonction puissance a déjà fait l'objet de nombreuses définitions et extensions. Voir le thème 52. Nous allons maintenant définir  $a^z$  pour  $a > 0$  et  $z \in \mathbb{C}$ .

Soit  $z = x + iy \in \mathbb{C}$ . L'exponentielle  $\exp(x + yi)$  est déjà définie en 15.66 ; il suffit donc maintenant de définir les notations  $e^z$  et  $a^z$  pour  $z \in \mathbb{C}$ .

$\left[ \begin{smallmatrix} \text{DEFooRBTDOoNLcWGj} \end{smallmatrix} \right]$

### Définition 18.5.

Pour le nombre  $e \in \mathbb{R}$  et le nombre imaginaire pur  $iy$  ( $y \in \mathbb{R}$ ), nous définissons

$$e^{iy} = \exp(iy) \quad (18.8)$$

où  $\exp$  est la série usuelle de la définition 15.66. Pour un nombre complexe général  $x + yi$  nous définissons

$$e^{x+iy} = e^x e^{iy}. \quad (18.9)$$

Et enfin, si  $a > 0$  et si  $z \in \mathbb{C}$  nous définissons

$$a^z = e^{z \ln(a)}, \quad (18.10)$$

la fonction logarithme ici étant celle  $\ln: ]0, \infty[ \rightarrow \mathbb{R}$  définie par la proposition 15.91.

Si  $z \in \mathbb{C}$  et si  $n \in \mathbb{Z}$ , la définition de  $z^n$  ne pose pas de problème, c'est la définition 1.333.

DefJilXoM

**18.6.**

Soit  $z = x + iy \in \mathbb{C}$ . L'exponentielle  $\exp(x + yi)$  est déjà définie en 15.66 ; elle est la fonction donnée par

$$\begin{aligned} \exp: \mathbb{C} &\rightarrow \mathbb{C} \\ z &\mapsto \sum_{n=0}^{\infty} \frac{z^n}{n!}. \end{aligned} \quad (18.11)$$

PROPooXEYFooIEaPvU

**Proposition 18.7.**

Le rayon de convergence<sup>2</sup> de la série exponentielle est infini.

*Démonstration.* L'exponentielle est la série de puissance dont les coefficients sont donnés par la suite  $(a_k) = 1/k!$ . Nous utilisons la formule de Hadamard de la proposition 15.39 :

$$\frac{1}{R} = \lim_{n \rightarrow \infty} \left| \frac{1/(n+1)!}{1/n!} \right| = \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!} = \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0. \quad (18.12)$$

Donc  $R = \infty$  par la partie de Hadamard qui dit que si on tombe sur  $1/R = 0$ , alors c'est  $R = \infty$ .  $\square$

PROPooWSDKooJREQGk

**Proposition 18.8.**

Pour tout  $z \in \mathbb{C}$  nous avons

$$\exp(z) = e^z. \quad (18.13)$$

PropdJjisy

**Proposition 18.9 ([599]).**

Quelques propriétés de l'exponentielle.

- (1) La fonction  $\exp$  est continue.
- (2) Nous avons la formule  $e^{z+w} = e^z e^w$  pour tout  $z, w \in \mathbb{C}$ .
- (3)  $(e^z)^{-1} = e^{-z}$
- (4)  $(\exp(z))^n = \exp(nz)$ .

ITEMooRLHCooJTuyKv

ITEMooIFYFooUniuKS

*Démonstration.* La proposition 18.7 nous enseigne que le rayon de convergence est infini. La fonction ainsi définie est alors continue par la proposition 12.393.

Les séries  $\exp(z)$  et  $\exp(w)$  ayant un rayon de convergence infini, nous pouvons utiliser le produit de Cauchy (théorème 15.27) :

$$e^z e^w = \sum_{n=0}^{\infty} \left( \sum_{i+j=n} \frac{z^i w^j}{i! j!} \right) \quad (18.14a)$$

$$= \sum_{n=0}^{\infty} \left( \sum_{i=0}^n \frac{z^i w^{n-i}}{i! (n-i)!} \right) \quad (18.14b)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i=0}^n \binom{n}{i} z^i w^{n-i} \quad (18.14c)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} (z + w)^n \quad (18.14d)$$

$$= \exp(z + w). \quad (18.14e)$$

Nous avons utilisé la formule du binôme (proposition 1.375).

Les autres propriétés énoncées sont des corollaires :

$$e^z e^{-z} = e^0 = 1. \quad (18.15)$$

 $\square$ 


---

2. Définition 15.12.

D'autres propriétés de l'exponentielle sur  $\mathbb{C}$ , entre autres l'holomorphie, sont données dans le théorème 26.108.

LEMooTDGKooWdpUTD

**Lemme 18.10** ([1]).

Soient  $a > 0$ ,  $z \in \mathbb{C}$  et  $n \in \mathbb{Z}$ . Alors

$$(a^z)^n = a^{nz}. \quad (18.16)$$

*Démonstration.* Il s'agit d'un calcul utilisant les propositions 18.9(4) et 18.8 :

$$(a^z)^n = (e^{z \ln(a)})^n \quad (18.17a)$$

$$= \exp(z \ln(a))^n \quad (18.17b)$$

$$= \exp(nz \ln(a)) \quad (18.17c)$$

$$= e^{nz \ln(a)} \quad (18.17d)$$

$$= a^{nz}. \quad (18.17e)$$

□

### 18.1.3 Formules de trigonométrie

Le lemme suivant est un premier pas pour le paramétrage du cercle dont nous parlerons dans la proposition 18.62(1).

LEMooHOYZooKQTsXW

**Lemme 18.11.**

Pour tout  $x \in \mathbb{R}$  nous avons :

(1)

$$e^{ix} = \cos(x) + i \sin(x) \quad (18.18)$$

EQooRVPJooTMwNTU

(2)  $|e^{ix}| = 1$ .

*Démonstration.* La définition de l'exponentielle sur  $\mathbb{C}$  est la définition 15.66. Cette définition fonctionne parce que  $\mathbb{C}$  est une algèbre normée, et que  $\mathbb{C}$  est un  $\mathbb{C}$ -module vérifiant l'inégalité  $|zz'| \leq |z||z'|$  (en l'occurrence, une égalité).

Nous remarquons que  $i^k$  vaut 1,  $i$ ,  $-1$ ,  $-i$ . Donc un terme sur deux est imaginaire pur et parmi ceux-là, un sur deux est positif. À bien y regarder, les termes imaginaires purs forment la série du sinus et ceux réels la série du cosinus.

Si vous aimez les formules,

$$e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} = \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{(2n+1)!} = \cos(y) + i \sin(y). \quad (18.19)$$

Nous avons utilisé le fait que  $i^{2n} = (-1)^n$  et  $i^{2n+1} = i(-1)^n$ .

□

NORMooXOVSoeEwjecU

**18.12** (Formule d'Euler).

En posant  $x = \pi$  dans (18.18) nous trouvons la soi-disant géniale formule d'Euler :

$$e^{i\pi} = -1. \quad (18.20)$$

Cette formule fait intervenir les nombres  $\pi$ ,  $e$ ,  $-1$  et  $i$ .

En fait non.

Le nombre  $e$  n'intervient pas dans la formule d'Euler. Vous vous souvenez de la définition (18.5) ? En fait  $e^z$  est juste un abus de notation pour  $\exp(z)$ .

Lorsque  $x$  est réel, nous avons la formule  $\exp(x) = e^x$  qui est un théorème tout à fait non trivial liant la fonction exponentielle et la fonction puissance. Il n'y a pas de tel résultat pour  $z$  complexe. Dans le cas de  $z$  complexe, la formule  $\exp(z) = e^z$  est juste une définition de la fonction puissance dans les complexes.

Bref, le nombre  $e$  n'intervient pas dans la formule d'Euler.

CORooWZFIOoDTCoRo

**Corolaire 18.13.**

Le complexe conjugué<sup>3</sup> de  $e^{ix}$  est  $e^{-ix}$ .

*Démonstration.* Vu le lemme 18.11, le complexe conjugué de  $z = e^{ix}$  est  $\bar{z} = \cos(x) - i \sin(x)$ . En utilisant (18.6c) nous avons également

$$\bar{z} = \cos(x) - i \sin(x) = \cos(-x) + i \sin(-x) = e^{-ix}. \quad (18.21)$$

□

LEMooJAWBooJGfZIL

**Lemme 18.14.**

Nous avons les formules d'addition d'angles<sup>4</sup> | SUBEQSooFSSMooHcYwRc

$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \quad (18.22a) \quad \text{EQooJYEMooQaQMib}$$

$$\sin(a+b) = \cos(a)\sin(b) + \sin(a)\cos(b) \quad (18.22b) \quad \text{EQooECAUpooQzckDv}$$

$$\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b) \quad (18.22c) \quad \text{EQooCVZAooQfocya}$$

$$\sin(a-b) = \sin(a)\cos(b) - \cos(a)\sin(b). \quad (18.22d)$$

pour tout  $a, b$  réels.

*Démonstration.* Nous utilisons la formule d'addition dans l'exponentielle, proposition (15.306) et la formule (18.18) avant de séparer les parties réelles et imaginaires :

$$e^{i(a+b)} = e^{ia}e^{ib} = \cos(a)\cos(b) - \sin(a)\sin(b) + i(\cos(a)\sin(b) + \sin(a)\cos(b)). \quad (18.23)$$

Cela est également égal à

$$\cos(a+b) + i \sin(a+b). \quad (18.24)$$

En identifiant les parties réelle et imaginaires, nous obtenons les formules (18.22a) et (18.22c) annoncées.

Pour la formule (18.22b), il suffit de se souvenir que  $\sin(-b) = -\sin(b)$  et  $\cos(-b) = \cos(b)$  (ces deux égalités sont immédiatement visibles sur les développements en série : l'un a uniquement des puissances paires et l'autre impaires) et d'écrire (18.22a) avec  $-b$  au lieu de  $b$ . □

CORooQZDQooWjMXTF

**Corolaire 18.15.**

Les formules suivantes pour les duplications d'angles s'ensuivent :

$$\cos(2a) = \cos^2(a) - \sin^2(a) \quad (18.25a)$$

$$\sin(2a) = 2 \cos(a) \sin(a). \quad (18.25b) \quad \text{SUBEQooLRJDooQuEux}$$

*Démonstration.* Poser  $b = a$  dans les relations du lemme 18.14. □

PROPooXGABooBASbsq

**Proposition 18.16** (Formules de linéarisation[600]).

Il existe des réels  $a_{n,k}$  tels que

$$\cos^n(x) = \sum_{k=0}^{\infty} a_{n,k} \cos(kx). \quad (18.26)$$

La somme est finie.

LEMooPQWWooMdPWUT

**Lemme 18.17.**

Un sous-groupe de  $(\mathbb{R}, +)$  est soit dense dans  $\mathbb{R}$  soit de la forme  $p\mathbb{Z}$  pour un certain réel  $p \neq 0$ .

3. Définition 1.597.

4. Rien ne nous empêche de donner ce nom à ces formules, mais seriez-vous capable de définir précisément le mot « angle » ?



*Démonstration.* Soit  $A$ , un sous-groupe de  $(\mathbb{R}, +)$  qui ne soit pas dense. Soit un intervalle  $]a, b[$  qui n'intersecte pas  $A$  (si vous voulez frimer, vous noterez ici que nous utilisons le fait que les intervalles ouverts forment une base de la topologie de  $\mathbb{R}$ <sup>5</sup>). Si  $d = |b - a|$ , l'ensemble  $A$  ne contient pas deux éléments séparés par strictement moins de  $d$ . Soit  $p$ , le plus petit élément strictement positif de  $A$ ; nous avons  $p \geq d$  (parce que  $0 \in A$  de toutes façons).

Puisque  $A$  est un groupe, nous avons  $p\mathbb{Z} \subset A$ .

Pour l'inclusion inverse, si  $x \in A$  est hors de  $p\mathbb{Z}$ , il existe un  $y \in p\mathbb{Z}$  avec  $|x - y| < p$ . Et donc le nombre  $|x - y|$  est dans  $A$  tout en étant plus petit que  $p$ . Contradiction.  $\square$

PROP00FRVC00KSgYUM

**Proposition-Définition 18.18** (Périodicité, le nombre  $\pi$ <sup>[601]</sup>).

Plusieurs choses à propos de la périodicité de la fonction  $\cos$ .

(1) La fonction  $\cos$  est périodique<sup>6</sup>.

ITEM00VPMW00BqidZG

(2) Un nombre  $T > 0$  est une période si et seulement si  $\cos(T) = 1$  et  $\sin(T) = 0$ .

Nous définissons le nombre  $\pi > 0$  comme étant la moitié de la période de la fonction  $\cos$  :

$$2\pi = \min\{T > 0 \text{ tel que } \cos(x + T) = \cos(x), \forall x \in \mathbb{R}\}. \quad (18.27)$$

*Démonstration.* Plusieurs étapes.

(1) **La fonction cosinus n'est pas toujours positive** Supposons d'abord que  $\cos(x) > 0$  pour tout  $x \in \mathbb{R}$ . Dans ce cas, la fonction  $\sin$  est strictement croissante. Mais les deux fonctions sont bornées par 1 du fait de la formule  $\cos^2(x) + \sin^2(x) = 1$ . La fonction  $\sin$  étant croissante et bornée, elle est convergente vers un réel par la proposition 12.58 :

$$\lim_{x \rightarrow \infty} \sin(x) = \ell \quad (18.28)$$

pour un certain  $\ell > 0$ . Avec ça nous avons aussi (pour cause de dérivée)  $\lim_{x \rightarrow \infty} \sin'(x) = 0$ , c'est-à-dire  $\lim_{x \rightarrow \infty} \cos(x) = 0$ . Mais vu que  $\cos^2(x) + \sin^2(x) = 1$ , nous en déduisons que  $\lim_{x \rightarrow \infty} \sin(x) = 1$ . Mézalor  $\lim_{x \rightarrow \infty} \cos'(x) = -1$ , ce qui veut dire que la fonction  $\cos$  n'est pas bornée. Cela est impossible. Nous en déduisons que  $\cos(x)$  n'est pas toujours positive.

(2) **Il existe  $T > 0$  tel que  $\cos(T) = 1$  et  $\sin(T) = 0$**  Par ce que nous venons de faire, il existe  $r > 0$  tel que  $\cos(r) = 0$ . Pour cette valeur, nous avons aussi obligatoirement  $\sin(r) = \pm 1$ . Nous avons aussi, en utilisant les formules (18.22),

$$\cos(2r) = \cos^2(r) - \sin^2(r) = -1 \quad (18.29a)$$

$$\sin(2r) = 2 \cos(r) \sin(r) = 0. \quad (18.29b)$$

et par conséquent

$$\cos(4r) = \cos^2(2r) - \sin^2(2r) = 1 \quad (18.30a)$$

$$\sin(4r) = 2 \cos(2r) \sin(2r) = 0. \quad (18.30b)$$

Donc  $T = 4r$  fonctionne.

(3) **Si  $T$  est une période** Nous entrons dans le vif de la preuve. Soit un  $T > 0$  tel que  $\cos(x + T) = \cos(x)$  pour tout  $x \in \mathbb{R}$ . Avec la formule d'addition d'angle dans le cosinus nous cherchons un  $T$  tel que

$$\cos(x + T) = \cos(x) \cos(T) - \sin(x) \sin(T) = \cos(x) \quad (18.31)$$

et donc tel que

$$\cos(x) (\cos(T) - 1) = \sin(x) \sin(T). \quad (18.32)$$

EQ00ELSA00LNtBnm

Nous dérivons cette équation :

$$-\sin(x) (\cos(T) - 1) = \cos(x) \sin(T). \quad (18.33)$$

EQ00CECF00LpxXaw

5. Proposition 7.142.

6. Définition 12.206.

Nous multiplions chacune des deux équations (18.32) et (18.33) par  $\sin(x)$  et  $\cos(x)$  pour obtenir les quatre relations suivantes :

$$\begin{aligned} \cos^2(x)(\cos(T) - 1) - \sin(x)\cos(x)\sin(T) &= 0 && \text{SUBEQooLGQXooLMLW} \\ -\sin(x)\cos(x)(\cos(T) - 1) - \cos^2(x)\sin(T) &= 0 && \text{SUBEQooCHTDooKwvYZF} \\ \sin(x)\cos(x)(\cos(T) - 1) - \sin^2(x)\sin(T) &= 0 && \text{SUBEQooEWPToooTLCUMf} \\ -\sin^2(x)(\cos(T) - 1) - \sin(x)\cos(x)\sin(T) &= 0 && \text{SUBEQooGBXTooCFekGJ} \end{aligned} \quad \begin{matrix} (18.34a) \\ (18.34b) \\ (18.34c) \\ (18.34d) \end{matrix}$$

En faisant (18.34a) moins (18.34d) nous trouvons  $\cos(T) = 1$ . Et en sommant (18.34b) avec (18.34c) nous avons  $-\sin(T) = 0$ .

- (4) Si  $T > 0$  est tel que  $\sin(T) = 0$  et  $\cos(T) = 1$  Alors les formules d'addition d'angle du lemme 18.14 donnent tout de suite

$$\cos(x + T) = \cos(x). \quad (18.35)$$

À ce niveau nous croyons avoir prouvé que  $\cos$  était périodique et que la période est donnée par

$$\min\{T > 0 \text{ tel que } \sin(T) = 0, \cos(T) = 1\}. \quad (18.36)$$

Or rien n'est moins sûr parce qu'il pourrait arriver que ce minimum n'existe pas, c'est-à-dire que l'infimum soit zéro. Autrement dit, il peut arriver que l'ensemble des périodes soit dense. Plus précisément, soit  $P \subset \mathbb{R}$  l'ensemble des périodes de  $\cos$ . C'est un sous-groupe de  $(\mathbb{R}, +)$  et le lemme 18.17 nous dit que  $P$  est soit dense dans  $\mathbb{R}$ , soit de la forme  $p\mathbb{Z}$  pour un  $p > 0$ .

Si  $P$  est dense, soient  $t \in \mathbb{R}$  et une suite  $(t_n)$  dans  $P$  telle que  $t_n \rightarrow t$ . Pour tout  $x$  et tout  $n$  nous avons

$$\cos(x + t_n) = \cos(x), \quad (18.37)$$

Comme la fonction cosinus est continue, nous pouvons passer à la limite et écrire  $\cos(x + t) = \cos(x)$ . Cela étant valable pour tout  $x$  et pour tout  $t$ , la fonction cosinus est constante. Or nous savons que ce n'est pas le cas, donc  $P$  n'est pas dense. Donc cosinus est périodique.  $\square$

PROPooKNLAooLwQHea

### Proposition 18.19.

La fonction  $\sin$  est périodique de période  $2\pi$  et

$$2\pi = \min\{T > 0 \text{ tel que } \sin(T) = 0, \cos(T) = 1\}. \quad (18.38)$$

*Démonstration.* La proposition 18.18 dit que  $\cos$  est périodique. Puisque  $\sin = -\cos'$  par le lemme 18.3, la fonction  $\sin$  est également périodique par le lemme 12.208. Si  $T$  est une période de  $\cos$ , alors  $T$  est une période de  $\sin$ .

Mais  $\sin' = \cos$ , de telle sorte que les périodes de  $\sin$  sont périodes de  $\cos$ . Bref,  $T$  est une période de  $\sin$  si et seulement si  $T$  est une période de  $\cos$ .  $\square$

### 18.20.

Notons que tout ceci ne nous donne pas la plus petite indication d'ordre de grandeur de la valeur de  $\pi$ . Cela peut encore être 0.1 autant que 500.

PROPooMWMDooJYIIs

### Proposition 18.21 ([601, 1]).

Des propriétés à la chaîne à propos des sinus, cosinus et de leurs périodes.

ITEMooRJZHooCXcKmM

- (1) Nous avons

$$2\pi = \min\{x > 0 \text{ tel que } \cos(x) = 1, \sin(x) = 0\}. \quad (18.39)$$

ITEMooTNHMOooUoJNC

- (2) Les fonctions  $\sin$  et  $\cos$  sont périodiques de période  $2\pi$ .

ITEMooSPZBooIQLUXh

- (3) Nous avons  $\cos(\pi) = -1$  et  $\sin(\pi) = 0$ .

- (4) Pour tout  $a \in \mathbb{R}$  nous avons

$$\cos(a + \pi) = -\cos(a) \quad (18.40a)$$

$$\sin(a + \pi) = -\sin(a). \quad (18.40b)$$

(5) Nous avons

$$\pi = \min\{x > 0 \text{ tel que } \cos(x) = -1, \sin(x) = 0\}. \quad (18.41)$$

(6) Nous avons

$$\begin{cases} \cos(\pi/2) = 0 \\ \sin(\pi/2) = 1. \end{cases} \quad (18.42a)$$

$$\begin{cases} \cos(\pi/2) = 0 \\ \sin(\pi/2) = 1. \end{cases} \quad (18.42b)$$

(7) Nous avons les formules

$$\begin{cases} \cos(x + \pi/2) = -\sin(x) \\ \sin(x + \pi/2) = \cos(x) \end{cases} \quad (18.43a)$$

$$\begin{cases} \cos(x + \pi/2) = -\sin(x) \\ \sin(x + \pi/2) = \cos(x) \end{cases} \quad (18.43b)$$

pour tout  $x \in \mathbb{R}$ .

(8) Nous avons

$$\frac{\pi}{2} = \min\{x > 0 \text{ tel que } \sin(x) = 1, \cos(x) = 0\}. \quad (18.44)$$

(9) Nous avons les valeurs

$$\begin{cases} \cos(\frac{3\pi}{2}) = 0 \\ \sin(\frac{3\pi}{2}) = -1. \end{cases} \quad (18.45a)$$

$$\begin{cases} \cos(\frac{3\pi}{2}) = 0 \\ \sin(\frac{3\pi}{2}) = -1. \end{cases} \quad (18.45b)$$

(10) Nous avons

$$\frac{\pi}{2} = \min\{x > 0 \text{ tel que } \cos(x) = 0\}. \quad (18.46)$$

(11) Pour tout  $x \in ]0, \frac{\pi}{2}[$ , nous avons  $\cos(x) > 0$  et  $\sin(x) > 0$ .

*Démonstration.* C'est parti.

(1) Le fond de la proposition 18.18 est que toutes les périodes  $T > 0$  vérifient  $\cos(T) = 1$  et  $\sin(T) = 0$ . La définition de  $2\pi$  est que c'est la plus petite période.

(2) En utilisant le fait que l'une est la dérivée de l'autre, si  $T$  est une période de  $\cos$  nous avons

$$\sin(x + T) = -\cos'(x + T) \quad (18.47a)$$

$$= -\lim_{\epsilon \rightarrow 0} \frac{\cos(x + T + \epsilon) - \cos(x + T)}{\epsilon} \quad (18.47b)$$

$$= -\lim_{\epsilon \rightarrow 0} \frac{\cos(x + \epsilon) - \cos(x)}{\epsilon} \quad (18.47c)$$

$$= -\cos'(x) \quad (18.47d)$$

$$= \sin(x). \quad (18.47e)$$

Nous déduisons que toute période de  $\cos$  est une période de  $\sin$ . De la même façon, nous pouvons prouver l'autre sens : toute période de  $\sin$  est une période de  $\cos$ .

(3) D'un côté nous avons

$$\cos(2\pi) = \cos^2(\pi) - \sin^2(\pi) = 1 \quad (18.48)$$

parce que  $\cos(2\pi) = \cos(0) = 1$ . Puisque  $\cos(\pi)$  et  $\sin(\pi)$  sont bornés par  $-1$  et  $1$ , nous devons avoir  $\sin(\pi) = 0$  et  $\cos(\pi) = \pm 1$ .

Mais d'un autre côté, le nombre  $2\pi$  est le plus petit  $T$  vérifiant  $\cos(T) = 1$ ,  $\sin(T) = 0$ . Donc, avoir  $\cos(\pi) = 1$  n'est pas possible. Nous concluons

$$\begin{cases} \cos(\pi) = -1 \\ \sin(\pi) = 0. \end{cases} \quad (18.49a)$$

$$\begin{cases} \cos(\pi) = -1 \\ \sin(\pi) = 0. \end{cases} \quad (18.49b)$$

- (4) Il s'agit d'utiliser les formules d'addition d'angles du lemme 18.14 pour calculer  $\cos(a + \pi)$  et  $\sin(a + \pi)$  en tenant compte du fait que  $\cos(\pi) = -1$  et  $\sin(\pi) = 0$ .
- (5) Soit  $a \in ]0, \pi[$  tel que  $\cos(a) = -1$  et  $\sin(a) = 0$ . Alors nous avons

$$\cos(a + \pi) = -\cos(\pi) = 1 \quad (18.50a)$$

$$\sin(a + \pi) = -\sin(\pi) = 0, \quad (18.50b)$$

ce qui donnerait  $a + \pi \in ]\pi, 2\pi[$  dont le cosinus est 1 et le sinus est zéro. Mais nous savons déjà que  $2\pi$  est le minimum pour cette propriété.

- (6) Nous avons

$$-1 = \cos(\pi) = \cos^2(\pi/2) - \sin^2(\pi/2), \quad (18.51)$$

donc  $\cos(\pi/2) = 0$  et  $\sin^2(\pi/2) = 1$ , ce qui donne  $\sin(\pi/2) = \pm 1$ .

Nous devons départager le  $\pm$ . Pour cela nous savons que  $\sin'(0) = \cos(0) = 1$  et que  $\sin(0) = 0$ , donc il existe  $\epsilon > 0$  tel que pour tout  $x \in ]0, \epsilon[$  nous avons  $0 < \cos(x) < 1$  et  $0 < \sin(x) < 1$  (nous avons aussi utilisé le lien entre dérivation et croissance de la proposition 12.209). Nous choisissons  $\epsilon$  plus petit que  $\pi/2$ .

Supposons que  $\sin(\pi/2) = -1$ . Le théorème des valeurs intermédiaires 9.95 dit qu'il existe  $x_0 \in ]\epsilon, \pi/2[$  tel que  $\sin(x_0) = 0$ . Pour cette valeur de  $x_0$  nous devons aussi avoir  $\cos(x_0) = \pm 1$ . Mais puisque  $2\pi$  est minimum pour avoir  $\cos = 1$  et  $\sin = 0$ , nous devons avoir  $\cos(x_0) = -1$ . Alors nous avons aussi

$$\cos(x_0 + \pi) = \cos(x_0)\cos(\pi) - \sin(x_0)\sin(\pi) = -\cos(x_0) = 1 \quad (18.52a)$$

$$\sin(x_0 + \pi) = \cos(x_0)\sin(\pi) + \sin(x_0)\cos(\pi) = \sin(x_0) = 0. \quad (18.52b)$$

Encore une fois par minimalité de  $2\pi$ , cela ne va pas. Conclusion :  $\sin(\pi/2) = 1$ .

- (7) Il s'agit encore d'utiliser les formules d'addition d'angle en tenant compte des valeurs remarquables  $\cos(\pi/2) = 0$  et  $\sin(\pi/2) = 1$ .
- (8) Supposons  $x_0 \in ]0, \pi/2[$  tel que  $\sin(x_0) = 1$  et  $\cos(x_0) = 0$ . En utilisant les formules (18.43) nous avons

$$\cos(x_0 + \pi/2) = -1 \quad (18.53a)$$

$$\sin(x_0 + \pi/2) = 0, \quad (18.53b)$$

avec  $x_0 + \pi/2 < \pi$ . Cela contredirait la minimalité de  $\pi$ .

- (9) Il s'agit d'utiliser les formules (18.43) :

$$\cos\left(\frac{3\pi}{2}\right) = \cos(\pi + \pi/2) = -\sin(\pi) = 0 \quad (18.54a)$$

$$\sin\left(\frac{3\pi}{2}\right) = \sin(\pi + \pi/2) = \cos(\pi) = -1. \quad (18.54b)$$

- (10) Si  $\cos(x_0) = 0$  alors  $\sin(x_0) = -1$  (parce que  $\sin(x_0) = 1$  est déjà exclu). Alors  $\cos(x_0 + \pi/2) = 1$  et  $\sin(x_0 + \pi/2) = 0$ , ce qui est également impossible.
- (11) La fonction cosinus est continue (proposition 18.1) et  $\cos(0) = 1$ . Le théorème des valeurs intermédiaires implique que si  $\cos(x) \leq 0$ , alors il existe  $t \in ]0, x]$  avec  $\cos(t) = 0$ . Cela n'est pas possible pour  $x < \pi/2$ , par le point (8).

Le cosinus est positif sur l'intervalle considéré et  $\sin'(x) = \cos(x)$ . Donc  $\sin(0) = 0$  et la dérivée est positive. La proposition 12.209 conclut que  $\sin$  est strictement croissante et donc, strictement positive.

□

LEMooFESYooBoiuol

**Lemme 18.22** (Positivité[1]).

À propos de positivité de la fonction cosinus.

$$(1) \cos(0) = 1$$

$$(2) \cos(x) > 0 \text{ pour } x \in [0, \pi/2[.$$

$$(3) \cos(\pi/2) = 0.$$

$$(4) \cos(x) < 0 \text{ pour } x \in ]\pi/2, 3\pi/2[$$

$$(5) \cos(3\pi/2) = 0.$$

$$(6) \cos(x) > 0 \text{ pour } x \in ]3\pi/2, 2\pi].$$

ITEMooIXSDooJyCQyb

ITEMooWJEVooGZykbO

ITEMooANEPooLGmYtc

ITEMooRDWJooZXWyfv

ITEMooFKPAooBNlvPU

ITEMooIDZGooBTDvDF

*Démonstration.* En plusieurs points.

(1) **Pour (1)** C'est déjà fait dans le lemme 18.2.

(2) **Pour (2)** C'est la proposition 18.21(11).

(3) **Pour (3)** C'est la proposition 18.21(6).

(4) **Pas d'annulation entre  $\pi/2$  et  $\pi$**  Nous montrons à présent que  $\cos$  ne s'annule pas entre  $\pi/2$  et  $\pi$ . Supposons que  $\cos(\frac{\pi}{2} + s) = 0$  avec  $s \in ]0, \pi/2[$ . Comme  $\cos(x)^2 + \sin(x)^2 = 1$  (lemme 18.4), nous avons

$$\begin{cases} \cos(\frac{\pi}{2} + s) = 0 \\ \sin(\frac{\pi}{2} + s) = \epsilon \end{cases} \quad (18.55a)$$

$$\quad (18.55b)$$

avec  $\epsilon = \pm 1$ . Utilisant trois fois la proposition 18.21(7) nous trouvons

$$\begin{cases} \cos(x + \frac{3\pi}{2}) = \sin(x) \\ \sin(x + \frac{3\pi}{2}) = -\cos(x) \end{cases} \quad (18.56a)$$

$$\quad (18.56b)$$

pour tout  $x$ . Nous appliquons cela à  $x = \frac{\pi}{2} + s$ , en nous souvenant que  $\cos(x + 2\pi) = \cos(x)$  et  $\sin(x + 2\pi) = \sin(x)$  (par 18.21(2)) :

$$\cos(s) = \cos(\frac{\pi}{2} + s + \frac{3\pi}{2}) = \sin(\frac{\pi}{2} + s) = \epsilon \quad (18.57)$$

et

$$\sin(s) = \sin(\frac{\pi}{2} + s + \frac{3\pi}{2}) = -\cos(\frac{\pi}{2} + s) = 0. \quad (18.58)$$

Si  $\epsilon = 1$ , nous avons une contradiction avec 18.21(1). Si  $\epsilon = -1$ , nous avons une contradiction avec 18.21(5).

Donc  $\cos(x) \neq 0$  pour  $x \in ]\frac{\pi}{2}, \pi]$ .

(5)  **$\cos(x) < 0$  sur  $]\pi/2, \pi]$**  Nous savons que  $\cos(\pi) = -1$  (18.21(3)). Étant donné que la fonction  $\cos$  est continue et qu'elle ne s'annule pas sur  $]\pi/2, \pi]$ , nous en déduisons qu'elle y est partout strictement négative.

(6) **Pour (4), (5), (6)** Il est directement visible sur le développement de définition que  $\cos(-x) = \cos(x)$ . Et comme  $\cos(x + 2\pi) = \cos(x)$ , nous avons

$$\cos(\pi + s) = \cos(-\pi - s) = \cos(\pi - s). \quad (18.59)$$

Donc toutes les valeurs (et tous les signes) de  $\cos(x)$  sur  $[\pi, 2\pi]$  peuvent être déduits de ceux sur  $[0, \pi]$ .

□

LEMooPARBooTXbbiB

### Lemme 18.23.

Soient  $x, y \in \mathbb{R}$ .

(1) Nous avons  $\cos(x) = \cos(y)$  si et seulement si

$$y \in \{x + 2k\pi\}_{k \in \mathbb{Z}} \cup \{-x + 2k\pi\}_{k \in \mathbb{Z}}. \quad (18.60)$$

(2) Nous avons  $\sin(x) = \sin(y)$  si et seulement si

$$y \in \{x + 2k\pi\}_{k \in \mathbb{Z}} \cup \{-x + (2k + 1)\pi\}_{k \in \mathbb{Z}}. \quad (18.61)$$

**Proposition 18.24.**

Les nombres  $x, y \in \mathbb{R}$  vérifient

$$\begin{cases} \cos(x) = \cos(y) \\ \sin(x) = \sin(y) \end{cases} \quad (18.62a)$$

$$\begin{cases} \cos(x) = \cos(y) \\ \sin(x) = \sin(y) \end{cases} \quad (18.62b)$$

si et seulement si il existe  $k \in \mathbb{Z}$  tel que  $y = x + 2k\pi$ .

*Démonstration.* En deux parties.

- (1)  $\Leftarrow$  Si  $y = x + 2k\pi$ , le résultat est correct parce que la proposition 18.21(2) dit que  $\sin$  et  $\cos$  sont périodiques de période  $2\pi$ .
- (2)  $\Rightarrow$  Supposons que  $x > y$ . Nous calculons  $\sin(x - y)$  et  $\cos(x - y)$  en utilisant les formules du lemme 18.14 et en tenant compte de (18.62). Cela donne  $\cos(x - y) = 1$  et  $\sin(x - y) = 0$ . La proposition 18.18(2) dit alors que  $x - y$  est une période de la fonction  $\cos$ . Or la période de  $\cos$  est  $2\pi$  (proposition 18.21(2)). Donc toutes les périodes de  $\cos$  sont les  $2k\pi$  avec  $k \in \mathbb{N}$  (lemme 12.207).

□

CORooTFMAooHDrRqi

**Corolaire 18.25.**

Des nombres  $x, y \in \mathbb{R}$  vérifient  $e^{ix} = e^{iy}$  si et seulement si il existe  $k \in \mathbb{Z}$  tel que  $y = x + 2k\pi$ .

*Démonstration.* Le lemme 18.11 donne  $e^{ix} = \cos(x) + i\sin(x)$ . Donc l'équation  $e^{ix} = e^{iy}$  revient au système (18.62) dont les solutions sont bien  $y = x + 2k\pi$ .

□

LEMooBIPFooQNiTqZ

**Lemme 18.26 ([1]).**

À propos de croissance et décroissance des fonctions trigonométriques.

- (1) Sur  $]0, \pi[$ , la fonction  $\cos$  est décroissante.
- (2) Sur  $]-\pi, 0[$ , la fonction  $\cos$  est croissante.

*Démonstration.* Nous savons que  $\cos' = -\sin$  par le lemme 18.3. La liaison entre dérivée et croissance est la proposition 12.209. Les signes de la fonction cosinus sont dans le lemme 18.22. Les signes de la fonction sinus peuvent être déduits de la proposition 18.21(7).

Vous avez tout en main.

□

Tout cela nous permet de calculer quelques valeurs remarquables de cosinus et sinus ainsi que d'écrire le tableau de variations de sinus et cosinus.

LEMooIGNPooPEctJy

**Lemme 18.27.**

Nous avons les valeurs remarquables

$$\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}. \quad (18.63)$$

*Démonstration.* La relation (18.25b) donne

$$0 = \cos(\pi/2) = \cos^2(\pi/4) - \sin^2(\pi/4). \quad (18.64)$$

Donc  $\cos^2(\pi/4) = \sin^2(\pi/4)$ . Mais puisque  $\sin(\pi/4)$  et  $\cos(\pi/4)$  sont positifs, ils sont égaux.

Nous avons aussi  $\sin^2(\pi/4) + \cos^2(\pi/4) = 1$ . Donc le nombre  $x = \cos(\pi/4) = \sin(\pi/4)$  vérifie l'équation  $2x^2 = 1$ , dont l'unique solution positive est  $x = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$ .

□

LEMooRMHAooDEAPMw

**Lemme 18.28.***Nous avons la valeur remarquable*

$$\cos\left(\frac{\pi}{3}\right) = \frac{1}{2}. \quad (18.65)$$

*Démonstration.* Il faut utiliser la formule (18.22a) avec  $\cos(\pi) = \cos(2\pi/3 + \pi/3)$  en sachant que  $\cos(\pi) = -1$ . Ensuite  $\cos(2\pi/3) = \cos(\pi/3 + \pi/3)$ . En décomposant ainsi, nous exprimons  $-1 = \cos(\pi)$  en termes de  $\cos(\pi/3)$  et de  $\sin(\pi/3)$ . En substituant  $\sin^2(\pi/3) = 1 - \cos^2(\pi/3)$  nous trouvons que le nombre  $\cos(\pi/3)$  vérifie l'équation

$$4x^3 - 3x + 1 = 0. \quad (18.66)$$

Croyez-le ou non, les solutions de cette équation sont  $x = -1$  et  $x = 1/2$ . Allez. Faisons comme si nous le savions pas. En tout cas, ces deux nombres sont des solutions, et nous avons la factorisation <sup>7</sup>

$$4x^3 - 3x + 1 = (2x - 1)^2(x + 1). \quad (18.67)$$

Donc  $1/2$  est de multiplicité 2 et  $-1$  de multiplicité 1. Le théorème 3.118 nous dit qu'il n'y a alors pas d'autres racines que ces deux-là <sup>8</sup>.

Nous en déduisons que la valeur de  $\cos(\pi/3)$  est soit  $1/2$  soit  $-1$ . La proposition 18.21(5) nous dit qu'il est impossible que  $\cos(\pi/3)$  soit égal à  $-1$  parce que  $\pi/3 < \pi$ . Donc  $\cos(\pi/3) = 1/2$  comme annoncé.  $\square$

**Remarque 18.29.**

Vous avez déjà sans doute vu la démonstration de  $\cos(30^\circ) = 1/2$  à partir de la figure 18.4. Il n'est pas possible de l'utiliser parce que cela n'est en réalité pas loin d'être la définition de l'angle entre deux droites.

Si vous voulez savoir la définition de l'angle entre deux droites, il faut passer par la définition 18.151, laquelle se base sur le lemme 18.135 qui, elle-même, se base sur la proposition 18.59.

Bref, à notre niveau, nous sommes encore loin de pouvoir faire des raisonnements trigonométriques sur base de géométrie dans les triangles.

PROPOoJFAGooYjRJcb

**Proposition 18.30.**

Pour tout  $x \in [0, \pi/4[$  nous avons  $\cos(x) > \sin(x)$ .

*Démonstration.* Nous posons  $f(x) = \cos(x) - \sin(x)$ . Elle vérifie  $f(0) = 1$ . En utilisant les dérivées du lemme 18.3, nous trouvons

$$f'(x) = -(\sin(x) + \cos(x)). \quad (18.68)$$

Mais sur  $]0, \pi/2[$  nous avons  $\cos(x) > 0$  et  $\sin(x) > 0$  (proposition 18.21(11)). Donc  $f$  est strictement décroissante. Elle ne peut donc passer qu'une seule fois par zéro. Le lemme 18.27 nous indique que  $f(\pi/4) = 0$ . Donc  $f(x) > 0$  sur  $[0, \pi/4[$ .  $\square$

PROPOoIBYEooClQtq

**Proposition 18.31.**

Quelques valeurs trigonométriques.

(1) Pour le sinus :

(1a)  $\sin(0) = 0$

(1b)  $\sin(\pi/6) = 1/2$

(1c)  $\sin(\pi/4) = \sqrt{2}/2$

(1d)  $\sin(\pi/3) = \sqrt{3}/2$

(1e)  $\sin(\pi/2) = 1$

(2) Pour le cosinus :

(2a)  $\cos(0) = 1$

(2b)  $\cos(\pi/6) = \sqrt{3}/2$

(2c)  $\cos(\pi/4) = \sqrt{2}/2$

(2d)  $\cos(\pi/3) = 1/2$

(2e)  $\cos(\pi/2) = 0$

(3) Pour la tangente :

(3a)  $\tan(0) = 0$

(3b)  $\tan(\pi/6) = \sqrt{3}/3$

(3c)  $\tan(\pi/4) = 1$

7. Factorisation d'un polynôme en sachant des racines, proposition 3.111.

8. Nous attirons votre attention sur le fait que cela n'est en aucun cas une trivialité.

$$(3d) \tan(\pi/3) = \sqrt{3}$$

$$(3e) \tan(\pi/2) \text{ est non défini.}$$

*Démonstration.* Plusieurs ont déjà été faites. Les autres ne seront pas démontrées dans l'ordre énoncé.

(1)  $\sin(0) = 0$  Substitution dans la définition (18.2).

(2)  $\sin(\pi/4) = \sqrt{2}/2$  C'est le lemme 18.27.

(3)  $\sin(\pi/3) = 1/\sqrt{2}$  Nous utilisons la formule  $\sin^2(x) + \cos^2(x) = 1$  avec  $x = \pi/3$ . Vu que  $\cos(\pi/3) = 1/2$  (lemme 18.28), cela donne  $\sin^2(\pi/3) = 3/4$ . Nous en déduisons que  $\sin(\pi/3)$  vaut  $\pm \frac{\sqrt{3}}{2}$ .

La proposition 18.21(5) nous dit que  $\sin$  est positive sur  $[0, \pi]$ . Donc c'est bien la possibilité  $\sqrt{3}/2$  qui est la bonne.

(4)  $\sin(\pi/6) = 1/2$  et  $\cos(\pi/6) = \sqrt{3}/2$  Nous partons de l'équation (18.25b) pour écrire

$$\sin(\pi/3) = 2 \cos(\pi/6) \sin(\pi/6). \quad (18.69)$$

Nous avons déjà vu que  $\sin(\pi/3) = \sqrt{3}/2$ . En posant  $x = \sin(\pi/6)$  nous avons également  $\cos(\pi/6) = \sqrt{1-x^2}$  parce que nous savons que la fonction cosinus est positive sur  $[0, \pi/2]$  (proposition 18.21(11)). Nous avons donc l'équation

$$\frac{\sqrt{3}}{2} = 2x\sqrt{1-x^2}. \quad (18.70)$$

Nous passons au carré et posons  $y = x^2$ . Après quelque manipulations,

$$16y^2 - 16y + 3 = 0. \quad (18.71)$$

Cela donne deux possibilités pour  $y$  :  $\frac{3}{4}$  et  $\frac{1}{4}$ . Puisque  $x > 0$ , nous pouvons simplement passer à la racine carrée :  $x = \sqrt{3}/2$  ou  $x = 1/2$ .

Notez que si nous avions posé  $x = \cos(\pi/6)$  au lieu de  $x = \sin(\pi/6)$ , nous aurions obtenu le même résultat. Donc  $\sin(\pi/6)$  et  $\cos(\pi/6)$  peuvent tous deux avoir les valeurs  $\sqrt{3}/2$  ou  $1/2$ . Cela fait 4 possibilités.

Étant donné que  $\sin^2(\pi/6) + \cos^2(\pi/6) = 1$ , les deux possibilités avec  $\sin(\pi/6) = \cos(\pi/6)$  sont exclues.

La proposition 18.30 nous dit aussi que  $\cos(\pi/6) > \sin(\pi/6)$ . Donc  $\cos(\pi/6) = \sqrt{3}/2$  et  $\sin(\pi/6) = 1/2$ .

(5)  $\sin(\pi/2) = 1$  C'est dans (18.42).

(6)  $\cos(0) = 1$  Substitution dans la définition.

(7)  $\cos(\pi/6) = \sqrt{3}/2$  Déjà fait avec le sinus de  $\pi/6$ .

(8)  $\cos(\pi/4) = \sqrt{2}/2$  Lemme 18.27.

(9)  $\cos(\pi/3) = 1/2$  Lemme 18.28.

(10)  $\cos(\pi/2) = 0$  Dans (18.42).

Toutes les valeurs pour la tangente s'obtiennent maintenant par la définition, en calculant  $\tan(x) = \frac{\sin(x)}{\cos(x)}$ .  $\square$

Voici un tableau qui rappelle les valeurs à retenir pour les fonctions sinus, cosinus et tangente.

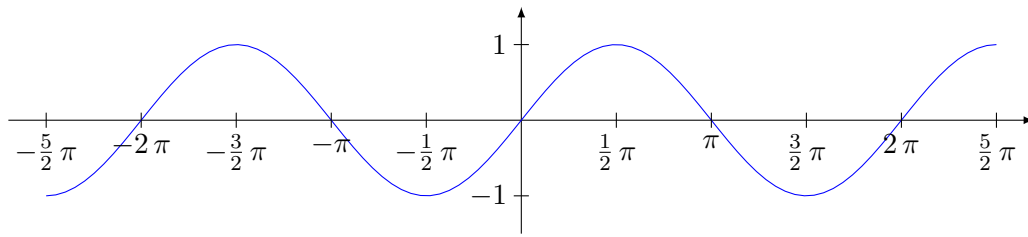
| $x$     | $\sin(x)$    | $\cos(x)$    | $\tan(x)$    |
|---------|--------------|--------------|--------------|
| 0       | 0            | 1            | 0            |
| $\pi/6$ | $1/2$        | $\sqrt{3}/2$ | $\sqrt{3}/3$ |
| $\pi/4$ | $\sqrt{2}/2$ | $\sqrt{2}/2$ | 1            |
| $\pi/3$ | $\sqrt{3}/2$ | $1/2$        | $\sqrt{3}$   |
| $\pi/2$ | 1            | 0            | N.D.         |

P GooIMQ FooTnBdU  
(18.72)

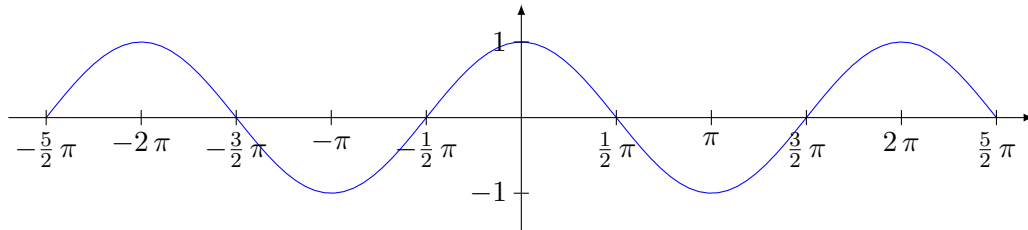
où « N.D. » signifie « non défini ».

Rappelons le graphe de la fonction sinus :





celui de la fonction cosinus :



developcosenpisur3

### Exemple 18.32.

Développer la fonction cos autour de  $x = \frac{\pi}{3}$ . Utiliser la valeur remarquable du lemme 18.28. Nous développons autour de  $h = 0$  la fonction  $\cos(\frac{\pi}{3} + h)$  :

$$\cos\left(\frac{\pi}{3} + h\right) \sim \cos\left(\frac{\pi}{3}\right) + h \cos'\left(\frac{\pi}{3}\right) + \frac{h^2}{2} \cos''\left(\frac{\pi}{3}\right) = \frac{1}{2} - \frac{\sqrt{3}}{2}h - \frac{1}{4}h^2. \quad (18.73)$$

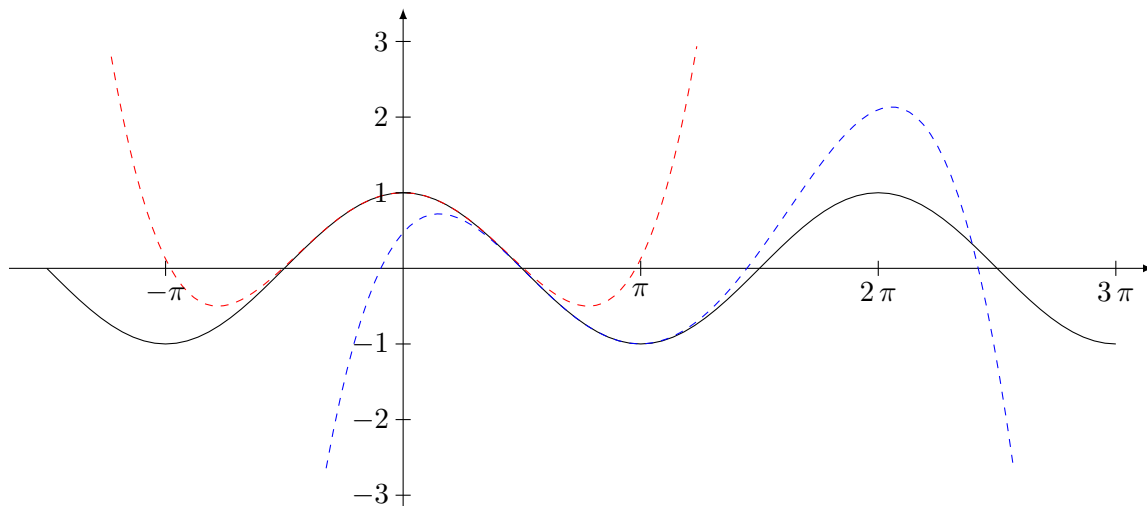
Il est aussi possible d'écrire cela en notant  $x = x_0 + h$ , c'est-à-dire en remplaçant  $h$  par  $x - \frac{\pi}{3}$  :

$$\cos(x) \sim \frac{1}{2} - \frac{\sqrt{3}}{2}\left(x - \frac{\pi}{3}\right) - \frac{1}{4}\left(x - \frac{\pi}{3}\right)^2. \quad (18.74)$$

△

### 18.33.

Voici un petit dessin pour donner une idée.



- (1) En noir le graphe de  $\cos(x)$ .
- (2) En rouge, le développement de  $\cos(x)$  à l'ordre 4 autour de  $x = 0$ .
- (3) En bleu, le développement de  $\cos(x)$  à l'ordre 4 autour de  $x = 3\pi/4$ .

## 18.2 Trucs et astuces de calcul d'intégrales

SECooKSOFooEVKDLh

Afin d'alléger le texte de calculs parfois un peu longs, nous regroupons ici les intégrales à une variable.

## 18.2.1 Quelques intégrales « usuelles »

ItemIntegrals

(1) L'intégrale

$$I = \int x \ln(x) dx = \frac{x^2}{2} \left( \ln(x) - \frac{1}{2} \right) \quad (18.75)$$

se fait par partie en posant

$$\begin{aligned} u &= \ln(x), & dv &= x dx \\ du &= \frac{1}{x} dx, & v &= \frac{x^2}{2}, \end{aligned} \quad (18.76)$$

et ensuite

$$I = \ln(x) \frac{x^2}{2} - \int \frac{x}{2} = \frac{x^2}{2} \left( \ln(x) - \frac{1}{2} \right). \quad (18.77)$$

(2) L'intégrale

$$I = \int x \ln(x^2) dx = x^2 \ln(x) - \frac{x^2}{2}. \quad (18.78)$$

En utilisant le fait que  $\ln(u^2) = 2 \ln(u)$ , nous retombons sur une intégrale du type (1) :

$$I = x^2 \ln(x) - \frac{x^2}{2}. \quad (18.79)$$

(3) L'intégrale

$$I = \int x \ln(1 + x^2) dx = \frac{1}{2} \ln(x^2 + 1)(x^2 + 1) - x^2 - \frac{1}{2} \quad (18.80)$$

EqTrucIntxlnxsqun

se traite en posant  $v = 1 + x^2$  de telle sorte à avoir  $dx = \frac{dv}{2x}$  et donc

$$I = \frac{1}{2} \ln(x^2 + 1)(x^2 + 1) - x^2 - \frac{1}{2}. \quad (18.81)$$

(4) L'intégrale

$$I = \int \cos(\theta) \sin(\theta) \ln \left( 1 + \frac{1}{\cos^2(\theta)} \right) d\theta \quad (18.82)$$

demande le changement de variable  $u = \cos(\theta)$ ,  $d\theta = -\frac{du}{\sin(\theta)}$ . Nous tombons sur l'intégrale

$$I = - \int u \ln \left( \frac{1 + u^2}{u^2} \right) = - \int u \ln(1 + u^2) + \int u \ln(u^2), \quad (18.83)$$

qui sont deux intégrales déjà faites. Nous trouvons

$$I = -\frac{1}{2} \ln \left( \frac{\sin^2(\theta) - 1}{\sin^2(\theta) - 2} \right) \sin^2(\theta) - \ln(\sin^2(\theta) - 2) + \frac{1}{2} \ln(\sin^2(\theta) - 1) \quad (18.84)$$

(5) L'intégrale

$$\int \frac{r^3}{1 + r^2} dr = \frac{r^2}{2} - \frac{1}{2} \ln(r^2 + 1). \quad (18.85)$$

commence par faire la division euclidienne de  $r^3$  par  $r^2 + 1$ ; ce que nous trouvons est  $r^3 = (r^2 + 1)r - r$ . Il reste à intégrer

$$\int \frac{r^3}{1 + r^2} dr = \int r dr - \int \frac{r}{1 + r^2} dr. \quad (18.86)$$

La fonction dans la seconde intégrale est  $\frac{r}{1+r^2} = \frac{1}{2} \frac{f'(r)}{f(r)}$  où  $f(r) = 1 + r^2$ , et donc  $\int \frac{r}{1+r^2} = \frac{1}{2} \ln(1 + r^2)$ . Au final,

$$I = \frac{1}{2} r^2 - \frac{1}{2} \ln(r^2 + 1). \quad (18.87)$$

(6) L'intégrale

$$I = \int \cos(\theta) \sin(\theta) d\theta = \frac{\sin^2(\theta)}{2} \quad (18.88)$$

se traite par le changement de variable  $u = \sin(\theta)$ ,  $du = \cos(\theta)d\theta$ , et donc

$$\int \cos(\theta) \sin(\theta) d\theta = \int u du = \frac{u^2}{2} = \frac{\sin^2(\theta)}{2}. \quad (18.89)$$

(7) L'intégrale

$$\int \sqrt{1+x^2} dx = \frac{x}{2} \sqrt{1+x^2} + \frac{1}{2} \operatorname{arcsinh}(x) \quad (18.90)$$

s'obtient en effectuant le changement de variable  $u = \sinh(\xi)$ .

(8) L'intégrale

$$\int \cos^2(x) \sin^2(x) dx = \frac{x}{8} - \frac{\sin(4x)}{32} \quad (18.91)$$

s'obtient à coups de formules de trigonométrie. D'abord, la formule (18.25b) permet de récrire la fonction à intégrer sous la forme

$$f(x) = \frac{1}{4} \sin^2(x). \quad (18.92)$$

Ensuite nous utilisons le fait que  $\sin^2(t) = (1 - \cos(2t))/2$  pour transformer la formule à intégrer en

$$f(x) = \frac{1 - \cos(4x)}{8}. \quad (18.93)$$

Cela s'intègre facilement en posant  $u = 4x$ , et le résultat est

$$\int f(x) dx = \frac{x}{8} - \frac{\sin(4x)}{32}. \quad (18.94)$$

(9) La fonction

$$\operatorname{sinc}(x) = \frac{\sin(x)}{x} \quad (18.95)$$

est le **sinus cardinal** de  $x$ . Nous allons montrer que

$$\int_0^\infty |\operatorname{sinc}(x)| dx = \infty. \quad (18.96)$$

D'abord nous avons

$$\int_{(n-1)\pi}^{n\pi} \frac{|\sin(t)|}{t} dt \geq \int_{(n-1)\pi}^{n\pi} \frac{|\sin(t)|}{n\pi} dt, \quad (18.97)$$

mais par périodicité,

$$\int_{(n-1)\pi}^{n\pi} |\sin(t)| dt = \int_0^\pi \sin(t) dt = 2. \quad (18.98)$$

Par conséquent

$$\int_0^{n\pi} |\operatorname{sinc}(t)| dt \geq \frac{2}{\pi} \sum_{k=1}^n \frac{1}{k}, \quad (18.99)$$

ce qui diverge lorsque  $n \rightarrow \infty$ .

(10) Les intégrales, pour  $\epsilon > 0$ ,

$$\int_0^\infty \cos(kx)e^{-\epsilon x} dx = \frac{\epsilon}{k^2 + \epsilon^2} \quad (18.100)$$

et

$$\int_0^\infty \sin(kx)e^{-\epsilon x} dx = \frac{k}{k^2 + \epsilon^2} \quad (18.101)$$

se calculent deux fois par partie. Nous posons

$$I = \int_0^\infty \cos(kx)e^{-\epsilon x} dx \quad (18.102a)$$

$$J = \int_0^\infty \sin(kx)e^{-\epsilon x} dx. \quad (18.102b)$$

L'intégrale  $I$  s'effectue par partie en posant  $u = \cos(kx)$  et  $v' = e^{-\epsilon x}$ . Un peu de calcul montre que

$$I = \frac{1}{\epsilon} - \frac{k}{\epsilon} J. \quad (18.103)$$

Par ailleurs l'intégrale  $J$  se fait également par partie pour obtenir

$$J = \frac{k}{\epsilon} I. \quad (18.104)$$

En résolvant pour  $I$  et  $J$  les deux équations déduites, nous trouvons

$$I = \frac{\epsilon}{k^2 + \epsilon^2} \quad (18.105a)$$

$$J = \frac{k}{k^2 + \epsilon^2}. \quad (18.105b)$$

### 18.2.2 Reforme un carré au dénominateur

subsecCarreDenoPar

Lorsqu'on a un second degré au dénominateur, le bon plan est de reformer un carré parfait. Par exemple :

$$x^2 + 2x + 2 = (x + 1)^2 + 1. \quad (18.106)$$

Ensuite, le changement de variable  $t = x + 1$  est pratique parce que cela donne  $t^2 + 1$  au dénominateur.

Cherchons

$$I = \int \frac{1-x}{x^2 + 2x + 2} dx = \int \frac{1-x}{(x+1)^2 + 1} dx = \int \frac{1-(t-1)}{t^2 + 1} \quad (18.107)$$

où nous avons fait le changement de variable  $t = x + 1$ ,  $dt = dx$ . L'intégrale se coupe maintenant en deux parties :

$$I = \int \frac{-t}{t^2 + 1} + \int \frac{2}{t^2 + 1}. \quad (18.108)$$

La seconde est dans les formulaires et vaut

$$2 \arctan(t) = 2 \arctan(x + 1), \quad (18.109)$$

tandis que la première est presque de la forme  $f'/f$  :

$$\int \frac{t}{t^2 + 1} = \frac{1}{2} \int \frac{2t}{t^2 + 1} = \frac{1}{2} \ln(t^2 + 1) = \frac{1}{2} \ln(u^2 + 2u + 2). \quad (18.110)$$

### 18.2.3 Décomposition en fractions simples

La décomposition en fractions simples décrite en 19.1.3 permet d'intégrer des fractions rationnelles. Elle peut parfois être évitée par la méthode de Rothstein-Trager que nous expliquerons dans 20.15.3.

### 18.3 Très modeste approximation de $\pi$

Nous sommes en droit de vouloir une valeur approchée de  $\pi$ .

LEMooJWSGooExmtDA

#### Lemme 18.34.

Nous avons l'encadrement

$$2\sqrt{2} < \pi < 4. \quad (18.111)$$

*Démonstration.* Grace au lemme 18.27 nous savons que la fonction  $\sin$  passe de 0 à  $\sqrt{2}/2$  sur un intervalle de taille  $\pi/4$  avec une dérivée majorée par 1. Par conséquent

$$\frac{\pi}{4} > \frac{\sqrt{2}}{2} \quad (18.112)$$

et donc<sup>9</sup>

$$\pi > 2\sqrt{2} \simeq 2.82 \quad (18.113)$$

De plus la fonction  $\sin$  passe de 0 à  $\sqrt{2}/2$  sur un intervalle de taille  $\pi/4$  avec une dérivée minorée par  $\sqrt{2}/2$ , donc

$$\frac{\pi}{4} < \frac{\sqrt{2}/2}{\sqrt{2}/2}, \quad (18.114)$$

ce qui donne

$$\pi < 4. \quad (18.115)$$

□

Pour avoir une meilleur approximation de  $\pi$ , nous pouvons remarquer que  $\pi \in ]2.82, 4[$ , et que cet intervalle est suffisamment petit pour ne pas recouvrir l'intervalle correspondant pour  $2\pi$ . L'équation  $\cos(x) = -1$  possède donc une unique solution dans cet intervalle (et cette solution est  $\pi$ ). Nous pouvons donc faire une dichotomie pour trouver la valeur de  $\pi$ , pourvu que nous ayons une façon d'évaluer des valeurs de  $\cos(x)$  de façon pas trop ridicule.

LEMooIECUooQOGYyN

#### Lemme 18.35.

Pour toute valeur de  $x \in \mathbb{R}$  on a  $|\sin(x)| \leq |x|$ .

*Démonstration.* Nous commençons par les valeurs  $x > 0$ . Considérons la fonction

$$f(x) = x - \sin(x). \quad (18.116)$$

La dérivée de  $f$  vaut

$$f'(x) = 1 - \cos(x) \geq 0, \quad (18.117)$$

De plus  $f(0) = 0$ . Vu que  $f$  est croissante, nous avons  $f(x) \geq 0$  pour tout  $x \geq 0$ .

(1) **Pour**  $0 \leq x \leq \pi$  Nous avons  $|x| = x$  et  $|\sin(x)| = \sin(x)$ , et ce que nous venons de dire à propos de  $f$  suffit pour conclure.

(2) **Pour**  $x \geq \pi$  En utilisant l'approximation de  $\pi$  du lemme 18.34, nous avons  $\pi > 2\sqrt{2} > 2$  et donc<sup>10</sup>

$$|x| - |\sin(x)| > 2 - 1 = 1 > 0. \quad (18.118)$$

(3) **Pour**  $x < 0$  Si  $x \leq 0$ , ce que nous venons de faire nous assure que  $\sin(-x) \leq -x$ . Par ailleurs, nous savons déjà par (18.6c) que  $\sin(-x) = -\sin(x)$ . Nous avons donc

$$|\sin(x)| = \sin(-x) \leq -x = |x|. \quad (18.119)$$

□

9. Sérieusement, êtes vous capables de trouver une approximation de  $\sqrt{2}$  en ne vous basant que sur des choses vues jusqu'ici ?

10. Ouais ; l'approximation  $\pi > 2$  nous suffit. On est large.

### 18.3.1 Les fonctions tangente et arc tangente

**Définition 18.36.**

La fonction *tangente* est :

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \quad (18.120)$$

où  $\sin$  et  $\cos$  sont de la définition 18.1.

La fonction tangente n'est pas définie sur les points de la forme  $x = \frac{\pi}{2} + k\pi$ ,  $k \in \mathbb{Z}$ . Une interprétation géométrique, qui justifie le nom, est donnée sur la figure 18.1.

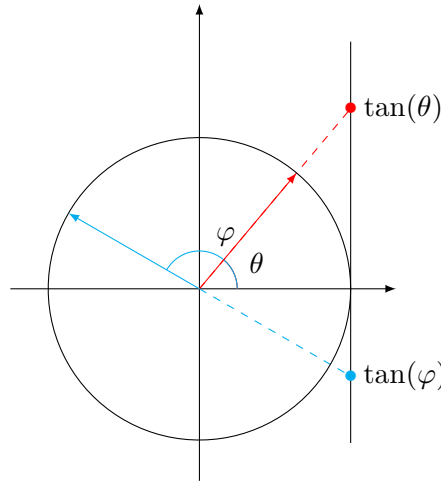


FIGURE 18.1: Interprétation géométrique de la fonction tangente. La tangente de l'angle  $\theta$  est positive (et un peu plus grande que 1) tandis que celle de l'angle  $\varphi$  est négative.

LEMooYAMFoooXCLXfr

**Lemme 18.37.**

La dérivée de la fonction tangente vaut

$$\tan'(x) = \tan^2(x) + 1, \quad (18.121)$$

*Démonstration.* Il suffit de suivre les règles de calcul comme la dérivation du quotient 12.194(5). □

PROPooTLHUoooPpqpE

**Proposition 18.38.**

La fonction

$$\begin{aligned} \tan: ]-\frac{\pi}{2}, \frac{\pi}{2}[ &\rightarrow \mathbb{R} \\ x &\mapsto \tan(x) \end{aligned} \quad (18.122)$$

est une bijection.

*Démonstration.* Le cosinus ne s'annulant pas sur l'intervalle donné, la fonction est bien définie. Nous avons

$$\lim_{x \rightarrow \pi/2^-} \tan(x) = +\infty \quad (18.123)$$

parce que la limite du sinus est 1 est celle du cosinus est zéro par les valeurs positives. Le même raisonnement donne la limite en  $-\pi/2$  qui vaut  $-\infty$ . Le théorème des valeurs intermédiaires<sup>11</sup> dit que la fonction tangente est alors surjective sur  $\mathbb{R}$ .

Par ailleurs le lemme 18.37 nous donne la dérivée de la tangente sous la forme  $\tan'(x) = \tan^2(x) + 1$ , ce qui nous donne une dérivée partout strictement positive, et donc une fonction strictement croissante et donc injective. □

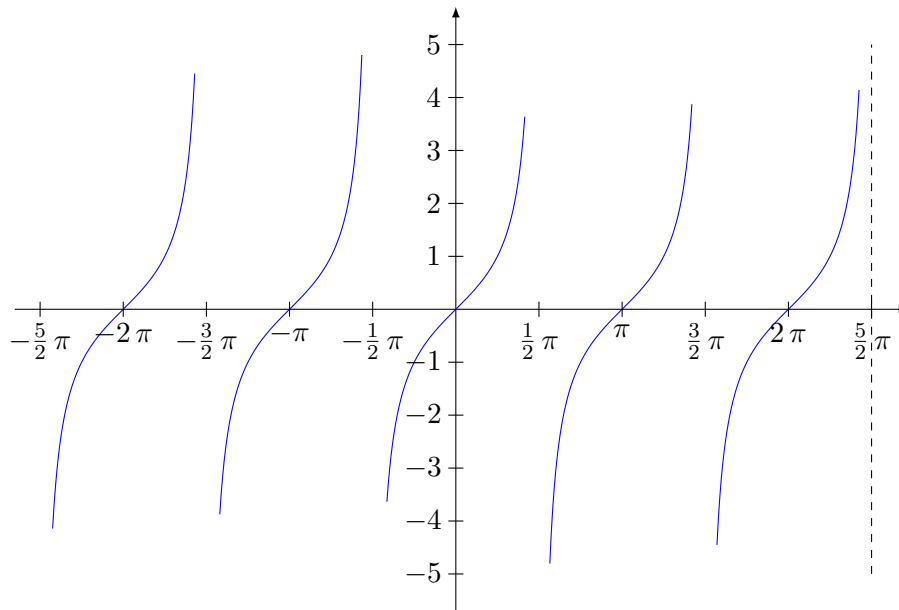


FIGURE 18.2: Le graphe de la fonction tangente.

LabelFigPVJooJDyNAg

Le graphe de la fonction tangente est sur la figure 18.2.

En ce qui concerne la bijection réciproque nous avons le théorème suivant.

THOooUSVGooOAnCvC

**Théorème 18.39.**

La fonction inverse de la tangente,

$$\begin{aligned} \arctan: \mathbb{R} &\rightarrow \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[ \\ x &\mapsto \arctan(x) \end{aligned} \quad (18.124)$$

nommée **arc tangente** est

(1) impaire et strictement croissante sur  $\mathbb{R}$ .

ITEMooMNHLoOVhIIB

(2) dérivable sur  $\mathbb{R}$  de dérivée

$$\arctan'(x) = \frac{1}{1+x^2}. \quad (18.125)$$

EQooGCHGooPlwYWt

*Démonstration.* Il est immédiatement visible sur son développement de définition (18.2) que la fonction sinus est impaire. Une vérification similaire montre que la fonction cosinus est paire. La fonction tangente est alors impaire et sa réciproque l'est tout autant.

La fonction arc tangente est également dérivable (donc continue) par la proposition 12.202 parce que la fonction tangente l'est. Notons qu'ici nous nous sommes restreint à  $]-\pi/2, \pi/2[$ . Sinon, le résultat est faux.

La formule proposée pour la dérivée provient également de la proposition 12.202 et de la dérivée de la tangente du lemme 18.37.  $\square$

LEMooHRDCooGtnyeQ

**Lemme 18.40.**

Nous avons les limites

$$(1) \lim_{x \rightarrow \infty} \arctan(x) = \frac{\pi}{2},$$

$$(2) \lim_{x \rightarrow -\infty} \arctan(x) = -\frac{\pi}{2}.$$

LEMooHJEKooARZMil

**Lemme 18.41.**

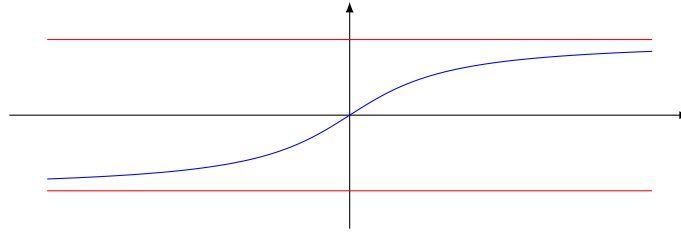
L'application  $\arctan: \mathbb{R} \rightarrow ]-\pi/2, \pi/2[$  est un homéomorphisme<sup>12</sup>.

11. Théorème 9.95.

12. Bijection continue, inversion continu définition 7.51.

**Exemple 18.42.**

Un homéomorphisme  $\phi: \mathbb{R} \rightarrow ]-1, 1[$  est donné par exemple par la fonction  $\phi(t) = \frac{2}{\pi} \arctan(t)$  dont le graphique est donné ci-dessous :



△

LEMooJKIUooEMMOrs

**Lemme 18.43.**

Nous avons la valeur remarquable

$$\arctan(1/\sqrt{3}) = \frac{\pi}{6}. \quad (18.126)$$

Le nombre  $\arctan(x_0)$  se calcule en cherchant l'angle  $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$  dont la tangente vaut  $x_0$ . Nous obtenons le tableau de valeurs suivant :

LEMooPQNCooDkEUyw

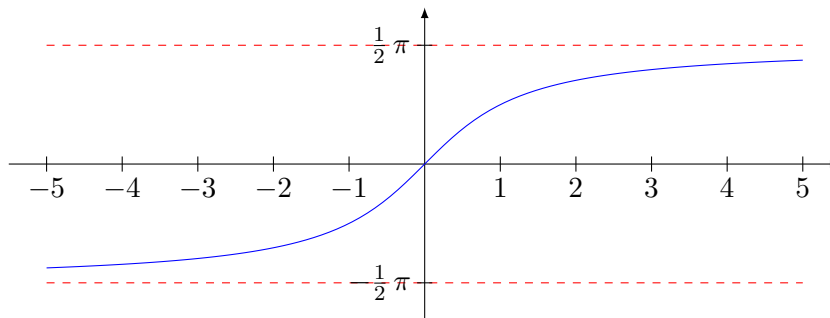
**Lemme 18.44.**

Quelques valeurs remarquables de l'arc tangente :

|              |   |                      |                 |                 |
|--------------|---|----------------------|-----------------|-----------------|
| $x$          | 0 | $\frac{1}{\sqrt{3}}$ | 1               | $\sqrt{3}$      |
| $\arctan(x)$ | 0 | $\frac{\pi}{6}$      | $\frac{\pi}{4}$ | $\frac{\pi}{3}$ |

(18.127)

En ce qui concerne la représentation graphique de la fonction  $x \mapsto \arctan(x)$ , elle s'obtient « en retournant » la partie entre  $-\frac{\pi}{2}$  et  $\frac{\pi}{2}$  du graphique de la fonction tangente :

**18.3.2 La fonction arc sinus**

Nous voulons étudier la fonction

$$\begin{aligned} \sin: \mathbb{R} &\rightarrow [-1, 1] \\ x &\mapsto \sin(x) \end{aligned} \quad (18.128)$$

et sa réciproque éventuelle.

La fonction sinus est continue sur  $\mathbb{R}$  mais n'est pas bijective : elle prend une infinité de fois chaque valeur de  $J = [-1, 1]$ . Pour définir une bijection réciproque de la fonction sinus en utilisant le théorème 12.53, nous devons donc choisir un intervalle à partir duquel la fonction sinus est monotone. Nous choisissons l'intervalle

$$I = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \quad (18.129)$$

La fonction

$$\begin{aligned} \sin: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] &\rightarrow [-1, 1] \\ x &\mapsto \sin(x) \end{aligned} \quad (18.130)$$

est une bijection croissante et continue. Nous avons donc le résultat suivant.



**Théorème 18.45** (Définition et propriétés de arc sinus).

Nous nommons **arc sinus** la bijection inverse de la fonction  $\sin: I \rightarrow J$ . La fonction

$$\begin{aligned} \arcsin: [-1, 1] &\rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ x &\mapsto \arcsin(x) \end{aligned} \quad (18.131)$$

ainsi définie est

- (1) continue et strictement croissante ;
- (2) impaire : pour tout  $x \in [-1, 1]$  nous avons  $\arcsin(-x) = -\arcsin(x)$ .

*Démonstration.* Nous prouvons le fait que  $\arcsin$  est impaire. Un élément de l'ensemble de définition de  $\arcsin$  est de la forme  $y = \sin(x)$  avec  $x \in [-\pi/2, \pi/2]$ . La relation (12.100) s'écrit dans notre cas

$$x = \arcsin(\sin(x)). \quad \text{EqVUWooUwVxVp} \quad (18.132)$$

Nous écrivons d'une part cette équation avec  $-x$  au lieu de  $x$  :

$$-x = \arcsin(\sin(-x)) = \arcsin(-\sin(x)) = \arcsin(-y); \quad \text{EqRLXooIwQySz} \quad (18.133)$$

et d'autre part nous multiplions (18.132) par  $-1$  :

$$-x = -\arcsin(\sin(x)) = -\arcsin(y). \quad \text{EqTGLooDeRYyT} \quad (18.134)$$

En égalisant les valeurs (18.133) et (18.134) nous trouvons

$$\arcsin(-y) = -\arcsin(y), \quad (18.135)$$

ce qui signifie que  $\arcsin$  est une fonction impaire.  $\square$

Notons que cette preuve repose sur le fait que tout élément de l'ensemble de définition de la fonction arc sinus peut être écrit sous la forme  $\sin(x)$  pour un certain  $x$ .

Si  $x_0 \in [-1, 1]$  est donné, calculer  $\arcsin(x_0)$  revient à trouver un angle  $\theta_0$  dans  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  pour lequel  $\sin(\theta_0) = x_0$ . Un tel angle sera forcément unique.

**Remarque 18.46.**

La définition de arc sinus découle du choix de l'intervalle  $I$ , qui est une convention. Il aurait été possible de faire un choix différent : pourriez-vous trouver la réciproque de la fonction sinus sur l'intervalle  $[\pi/2, 3\pi/2]$  ? Le mieux est de l'écrire comme une translatée de arc sinus, en utilisant le fait que sinus est une fonction périodique.

**Lemme 18.47.**

Nous avons  $\arcsin(1) = \frac{\pi}{2}$ .

*Démonstration.* Pour calculer  $\arcsin(1)$ , il faut chercher un angle entre  $-\frac{\pi}{2}$  et  $\frac{\pi}{2}$  ayant 1 pour sinus : résoudre  $\sin(\theta) = 1$ . La solution est  $\theta = \frac{\pi}{2}$  et nous avons donc  $\arcsin(1) = \frac{\pi}{2}$ .  $\square$

À l'aide des valeurs remarquables de la fonction sinus nous obtenons le tableau suivant de valeurs remarquables pour l'arc sinus.

|              |   |                 |                      |                      |                 |
|--------------|---|-----------------|----------------------|----------------------|-----------------|
| $x$          | 0 | $\frac{1}{2}$   | $\frac{\sqrt{2}}{2}$ | $\frac{\sqrt{3}}{2}$ | 1               |
| $\arcsin(x)$ | 0 | $\frac{\pi}{6}$ | $\frac{\pi}{4}$      | $\frac{\pi}{3}$      | $\frac{\pi}{2}$ |

Les autres valeurs remarquables peuvent être déduites du fait que l'arc sinus est une fonction impaire.

En ce qui concerne la dérivabilité de la fonction arc sinus, en application de la proposition 12.202 elle est dérivable en tout  $y = \sin(x)$  tel que  $\sin'(x) \neq 0$ , c'est-à-dire tel que  $\cos(x) \neq 0$ . Or  $\cos(x) = 0$  pour  $x = \pm\frac{\pi}{2}$ , ce qui correspond à  $y = \sin(\pm\frac{\pi}{2}) = \pm 1$ . La fonction arc sinus est donc dérivable sur  $] -1, 1[$ . Nous avons donc la propriété suivante pour la dérivabilité.

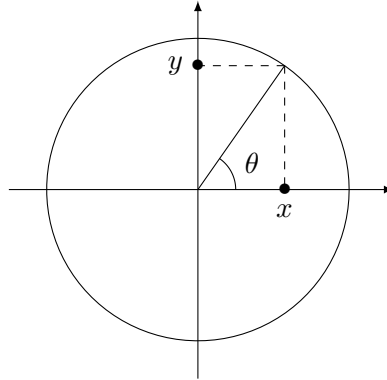
**Proposition 18.48.**

La fonction arc sinus est continue sur  $[-1, 1]$  et dérivable sur  $] -1, 1[$ . Pour tout  $y \in ] -1, 1[$ , la dérivée est donnée par la formule (12.529), qui dans ce cas s'écrit

$$\arcsin'(y) = \frac{1}{\cos(\arcsin(y))} = \frac{1}{\sqrt{1-y^2}}. \quad (18.136)$$

La dernière égalité viens du fait que si  $x = \arcsin(y)$  alors  $y = \sin(x)$  et  $\cos(x) = \sqrt{1 - \sin^2(x)} = \sqrt{1 - y^2}$ .

Pour comprendre la dernière égalité, remarquer que dans le dessin suivant,  $\theta = \arcsin(y)$ , donc  $y = \sin(\theta)$ , et  $x = \cos(\theta)$ .



Notons enfin que le graphe de la fonction arc sinus est donné à la figure 18.3.

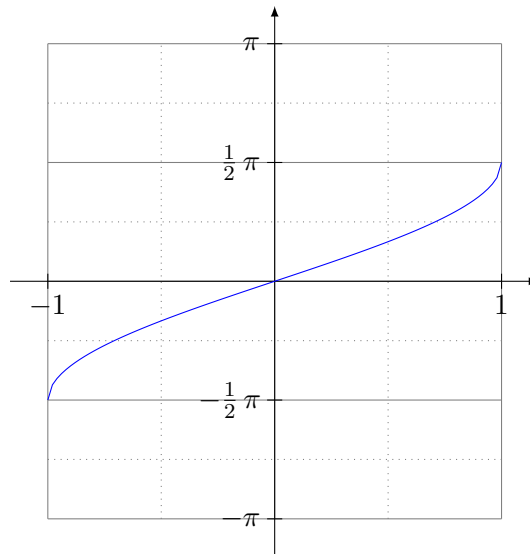


FIGURE 18.3: Le graphe de la fonction  $x \mapsto \arcsin(x)$

LabelFigFGRooDhFkch

### 18.3.3 La fonction arc cosinus

Nous voulons étudier la fonction

$$\cos: \mathbb{R} \rightarrow [-1, 1] \quad (18.137)$$

et son éventuelle réciproque. Encore une fois il n'est pas possible d'en prendre la réciproque globale parce que ce n'est pas une bijection ; ne fut-ce que parce qu'elle est périodique (proposition 18.18). Nous choisissons de considérer l'intervalle  $[0, \pi]$  sur lequel la fonction cosinus est continue et strictement monotone décroissante.

Nous avons alors le résultat suivant :

PROPoOZOZHooSMoYQD

**Proposition-Définition 18.49.**

Pour définir la fonction arc cosinus.

ITEMooMJWZooNHgQox

(1) La fonction

$$\cos: [0, \pi] \rightarrow [-1, 1] \tag{18.138}$$

est une bijection continue strictement décroissante.

(2) Sa bijection réciproque est la fonction

$$\arccos: [-1, 1] \rightarrow [0, \pi] \tag{18.139}$$

nommée **arc cosinus**.

(3) La fonction arc cosinus est continue, strictement décroissante.

(4) Elle est dérivable et pour tout  $y \in ]-1, 1[$ , sa dérivée est donnée par

$$\arccos'(y) = \frac{1}{-\sin(\arccos(y))} = \frac{-1}{\sqrt{1-y^2}}. \tag{18.140}$$

*Démonstration.* La fonction cosinus est continue et même de classe  $C^\infty$  par la proposition 18.1. Elle est strictement décroissante parce que sa dérivée  $(-\sin)$  y est strictement négative (strictement, à l'intérieur du domaine).

Le fait que arc cosinus soit une bijection continue strictement monotone est dans le théorème de la bijection 12.53. La dérivabilité et la formule sont de la proposition 12.202.  $\square$

Pour  $y_0 \in [-1, 1]$ , trouver la valeur de  $\arccos(y_0)$  revient à résoudre l'équation  $\cos(x_0) = y_0$ . Cela nous permet de construire une tableau de valeurs :

|              |       |                       |                       |                  |                  |                 |                      |                      |     |
|--------------|-------|-----------------------|-----------------------|------------------|------------------|-----------------|----------------------|----------------------|-----|
| $x$          | $-1$  | $-\frac{\sqrt{3}}{2}$ | $-\frac{\sqrt{2}}{2}$ | $-\frac{1}{2}$   | $0$              | $\frac{1}{2}$   | $\frac{\sqrt{2}}{2}$ | $\frac{\sqrt{3}}{2}$ | $1$ |
| $\arccos(x)$ | $\pi$ | $\frac{5\pi}{6}$      | $\frac{3}{4}\pi$      | $\frac{2}{3}\pi$ | $\frac{1}{2}\pi$ | $\frac{\pi}{3}$ | $\frac{1}{4}\pi$     | $\frac{1}{6}\pi$     | $0$ |

**Remarque 18.50.**

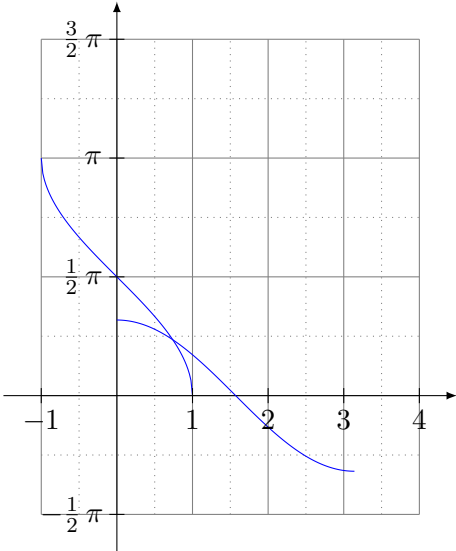
Certes la fonction cosinus est paire (vue sur  $\mathbb{R}$ ), mais la fonction arc cosinus ne l'est pas car elle est une bijection entre  $[-1, 1]$  et  $[0, \pi]$ .

**Exemple 18.51.**

Cherchons  $\arccos(\frac{1}{2})$ . Il faut trouver un angle  $\theta \in [0, \pi]$  tel que  $\cos(\theta) = \frac{1}{2}$ . La solution est  $\theta = \frac{\pi}{3}$ . Donc  $\arccos(\frac{1}{2}) = \frac{\pi}{3}$ .

Il n'est cependant pas immédiat d'en déduire la valeur de  $\arccos(-\frac{1}{2})$ . En effet  $\theta = \arccos(-\frac{1}{2})$  si et seulement si  $\cos(\theta) = -\frac{1}{2}$  avec  $\theta \in [0, \pi]$ . La solution est  $\theta = \frac{2\pi}{3}$ .  $\triangle$

En ce qui concerne la représentation graphique, il suffit de tracer la fonction cosinus entre 0 et  $\pi$  puis de prendre le symétrique par rapport à la droite  $y = x$ .



### 18.3.4 Une meilleure approximation de $\pi$

Nous avons laissé le nombre  $\pi$  avec l'approximation assez minable de  $2\sqrt{2} < \pi < 4$  en le lemme 18.34. Nous pouvons maintenant faire nettement mieux.

Le lemme 18.43 donne

$$\arctan(1/\sqrt{3}) = \pi/6 \quad (18.141)$$

et l'idée est de donner un développement de  $\arctan$  autour de zéro, de l'évaluer en  $1/\sqrt{3}$  et d'égaliser le résultat à  $\pi/6$ . Tout cela donne lieu à des calculs peut-être fastidieux, mais comme un gars l'a fait dès l'an 1424[602] pour trouver 16 décimales correctes, nous faisons comme si c'était facile.

Pour trouver le développement en série de Taylor (théorème 12.498) de arc tangente autour de  $x = 0$ , il faut partir de la formule (18.125) et sans doute pas mal calculer et faire une récurrence<sup>13</sup>. Le résultat est :

$$\arctan(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1}, \quad (18.142)$$

valable pour  $x \in ]-1, 1[$ . Avec cela nous avons

$$\arctan\left(\frac{1}{\sqrt{3}}\right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)3^k} \times \frac{1}{\sqrt{3}} = \frac{\pi}{6}, \quad (18.143)$$

et donc

$$\pi = \frac{6}{\sqrt{3}} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)3^k}. \quad (18.144)$$

Pour donner une idée du fait que ça fonctionne pas mal, voici le calcul pour quelques termes :

```

1 SageMath version 7.3, Release Date: 2016-08-04
2 Type "notebook()" for the browser-based notebook interface.
3 Type "help()" for help.
4
5
6 sage: n(taylor(arctan(x),x,0,5)(1/sqrt(3)))*6
7 3.15618147156995
8 sage: n(taylor(arctan(x),x,0,10)(1/sqrt(3)))*6
9 3.14260474566308
10 sage: n(taylor(arctan(x),x,0,20)(1/sqrt(3)))*6-pi
11 -2.14265171338823e-6
12 sage: n(taylor(arctan(x),x,0,58)(1/sqrt(3)))*6-pi
13 8.88178419700125e-16

```

tex/sage/sageSnip012.sage

Calculer 5 termes donne déjà 3.15. Et on est à  $10^{-6}$  de la bonne réponse avec 20 termes. Et avec 58 termes, on n'est à  $10^{-16}$ .

#### ⚠ Avertissement/question au lecteur !! 18.52

Pour bien faire, il faudrait étudier le reste et donner un encadrement.

### 18.3.5 Angle entre deux vecteurs

#### Proposition-Définition 18.53.

Soient des vecteurs  $X, Y \in \mathbb{R}^2$ . Il existe un unique  $\theta \in [0, \pi]$  tel que

$$\cos(\theta) = \frac{X \cdot Y}{\|X\| \|Y\|}. \quad (18.145)$$

Ce réel est appelé **angle** entre  $X$  et  $Y$ .

13. Je n'ai pas fait le calcul, merci de me faire savoir si il y a une astuce.

*Démonstration.* Si  $a$  et  $b$  sont des réels, l'inégalité  $|a| \leq b$  peut se développer en une double inégalité

$$-b \leq a \leq b. \quad (18.146)$$

L'inégalité de Cauchy-Schwarz (10.2) devient alors

$$-\|X\|\|Y\| \leq X \cdot Y \leq \|X\|\|Y\|. \quad (18.147)$$

Si  $X \neq 0$  et  $Y \neq 0$ , nous en déduisons

$$-1 \leq \frac{X \cdot Y}{\|X\|\|Y\|} \leq 1. \quad (18.148)$$

Il existe donc par la proposition 18.49 un angle  $\theta \in [0, \pi]$  tel que

$$\cos(\theta) = \frac{X \cdot Y}{\|X\|\|Y\|}. \quad (18.149)$$

□

### 18.54.

Certains n'hésitent pas à écrire la formule

$$X \cdot Y = \|X\|\|Y\| \cos(\theta). \quad \text{eqPropCosThet} \quad (18.150)$$

comme une définition du produit scalaire. C'est ce qui arrive lorsqu'on définit les fonctions trigonométriques à partir de relations dans les triangles rectangles.

Nous pouvons prouver simplement que  $\sin(30^\circ) = \frac{1}{2}$  et  $\cos(30^\circ) = \frac{\sqrt{3}}{2}$  en s'inspirant de la figure 18.4.

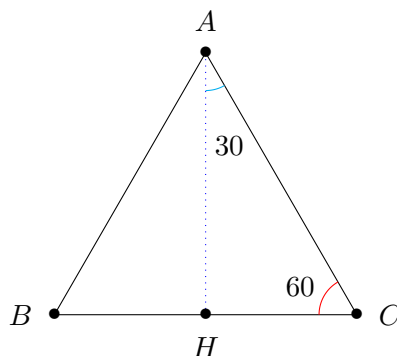


FIGURE 18.4: Un triangle équilatéral de côté 1.

LabelFigGVDJooYzMxLW

### 18.3.6 Aire du parallélogramme

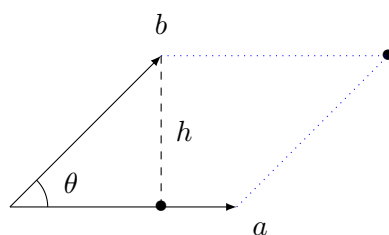


FIGURE 18.5: Calculer l'aire d'un parallélogramme.

LabelFigBNHLooLDxdPA

RemaAireParalProdVect

**Remarque 18.55.**

Le nombre  $\|a\|\|b\|\sin(\theta)$  est l'aire du parallélogramme<sup>14</sup> formé par les vecteurs  $a$  et  $b$ , comme cela se voit sur la figure 18.5. Un vrai calcul avec une intégrale sera effectué dans la proposition 20.26.

PropNormeProdVectoabsint

**Proposition 18.56.**

Nous avons

$$\|a \times b\| = \|a\|\|b\|\sin(\theta) \quad (18.151)$$

où  $\theta \in [0, \pi]$  est l'angle formé par  $a$  et  $b$ .

*Démonstration.* En utilisant la décomposition du produit vectoriel<sup>15</sup>, nous avons

$$\begin{aligned} \|a \times b\|^2 &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2 \\ &= (a_2b_3 - b_2a_3)^2 + (a_1b_3 - a_3b_1)^2 + (a_1b_2 - a_2b_1)^2 \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\ &= \|a\|^2\|b\|^2 - (a \cdot b)^2 \\ &= \|a\|^2\|b\|^2 - \|a\|^2\|b\|^2\cos^2(\theta) \\ &= \|a\|^2\|b\|^2(1 - \cos^2(\theta)) \\ &= \|a\|^2\|b\|^2\sin^2(\theta). \end{aligned} \quad (18.152)$$

D'où le résultat. Nous avons utilisé la formule de la définition (18.53) donnant l'angle en fonction du produit scalaire.  $\square$

NORMooWWOKooWzScnZ

**18.57.**

Si les vecteurs  $a$ ,  $b$  et  $c$  ne sont pas coplanaires, alors la valeur absolue du produit mixte (voir équation (10.125))  $a \cdot (b \times c)$  donne le volume du parallélépipède construit sur les vecteurs  $a$ ,  $b$  et  $c$ .

En effet si  $\varphi$  est l'angle entre  $b \times c$  et  $a$ , alors la hauteur du parallélépipède vaut  $\|a\|\cos(\varphi)$  parce que la direction verticale est donnée par  $b \times c$ , et la hauteur est alors la « composante verticale » de  $a$ . Par conséquent, étant donné que  $\|b \times c\|$  est l'aire de la base, le volume du parallélépipède vaut<sup>16</sup>

$$V = \|b \times c\|\|a\|\cos(\varphi). \quad (18.153)$$

Or cette formule est le produit scalaire de  $a$  par  $b \times c$ ; ce dernier étant donné par le déterminant de la matrice formée des composantes de  $a$ ,  $b$  et  $c$  grâce à la formule (10.125).

La valeur absolue du déterminant

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \quad (18.154) \quad \text{EgDeratb}$$

est l'aire du parallélogramme déterminé par les vecteurs  $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$  et  $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ . En effet, d'après la remarque 18.55, l'aire de ce parallélogramme est donnée par la norme du produit vectoriel

$$\begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} = \begin{vmatrix} e_x & e_y & e_z \\ a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} e_z, \quad (18.155)$$

donc la norme  $\|a \times b\|$  est bien donnée par la valeur absolue du déterminant (18.154).

14. Définition de ce qu'est une aire : 20.21. Preuve dans le cas d'un parallélogramme : 20.26.

15. Directement de la définition 10.36.

16. Le calcul de ce volume mériterait une certaine réflexion, surtout à partir du moment où nous avons décidé de définir les fonctions trigonométriques à partir de leur développement (définition 18.1).

## 18.4 Paramétrisation du cercle

Nous allons parler de paramétrisation du cercle. L'ensemble  $S^1$  sera vu tantôt comme le cercle dans  $\mathbb{R}^2$ , tantôt comme le cercle dans  $\mathbb{C}$ . Nous n'allons pas pousser le vice jusqu'à écrire explicitement les isomorphismes lorsque nous passons d'une représentation à l'autre. Parmi les identifications que nous allons faire sans ménagement, il y a l'identification entre les applications

$$\begin{aligned}\gamma: [0, 2\pi[ &\rightarrow \mathbb{R}^2 \\ t &\mapsto (\cos(t), \sin(t))\end{aligned}\tag{18.156}$$

et

$$\begin{aligned}\varphi: [0, 2\pi[ &\rightarrow \mathbb{C} \\ t &\mapsto e^{it}.\end{aligned}\tag{18.157}$$

C'est évidemment la formule  $e^{it} = \cos(t) + i\sin(t)$  (lemme 18.11) qui permet de transformer  $\gamma$  en  $\varphi$  et inversement. De plus  $\mathbb{R}^2$  et  $\mathbb{C}$  sont isomorphes en tant qu'espaces vectoriels normés (et aussi donc topologiques).

Nous allons prouver suffisamment de résultats à propos de ces deux applications pour pouvoir dire qu'elles paramétrisent le cercle et écrire des égalités du type

$$\int_{S^1} f = \int_0^{2\pi} f \circ \varphi.\tag{18.158}$$

## 18.5 Cercle trigonométriques

PROPOSWZFGooMVLtFz

**Proposition 18.58** ([603]).

Soient des fonctions  $f, g: I \rightarrow \mathbb{R}$  de classe  $C^1$  sur l'ouvert  $I$  de  $\mathbb{R}$  telles que  $f^2 + g^2 = 1$ . Soient  $t_0 \in I$  et  $\theta_0$  tel que  $f(t_0) = \cos(\theta_0)$  et  $g(t_0) = \sin(\theta_0)$ .

Alors il existe une unique fonction continue  $\theta: I \rightarrow \mathbb{R}$  telle que

$$\begin{cases} \theta(t_0) = \theta_0 \\ f = \cos \circ \theta \\ g = \sin \circ \theta. \end{cases}\tag{18.159a}$$

$$\tag{18.159b}$$

$$\tag{18.159c}$$

*Démonstration.* Nous commençons par l'existence, en passant par les nombres complexes. Soit  $h: I \rightarrow \mathbb{C}$  définie par  $h = f + ig$ . Nous avons  $h\bar{h} = 1$  et nous définissons

$$\theta(t) = \theta_0 - i \int_{t_0}^t h'(s) \overline{h(s)} ds.\tag{18.160}$$

Cette intégrale existe pour tout  $t$  parce que les fonctions  $f$  et  $g$  étant continues, elles sont bornées sur le compact  $[t_0, t]$ . De plus  $\theta$  est une fonction continue parce que c'est une primitive (proposition 14.287)<sup>17</sup>.

La dérivée de  $\theta$  est la fonction  $s \mapsto -ih'(s)\overline{h(s)}$ .

Utilisant la formule du lemme 18.11 sur la forme trigonométrique des nombres complexes, nous calculons :

$$\frac{d}{dt} [he^{-i\theta}]_{t=0} = e^{-i\theta} (h' - h\theta') = e^{-i\theta} (h' - ih(-i)h'\bar{h}) = 0.\tag{18.161}$$

Par conséquent il existe  $c \in \mathbb{C}$  tel que  $he^{-i\theta} = c$ . Mais  $h(t_0) = f(t_0) + ig(t_0) = \cos(\theta_0) + i\sin(\theta_0) = e^{i\theta_0}$ , du coup

$$h(t_0)e^{-i\theta(t_0)} = c\tag{18.162}$$

donne immédiatement  $c = 1$ , ou encore  $e^{i\theta(t)} = h(t)$ , c'est-à-dire que

$$f + ig = \cos \circ \theta + i \sin \circ \theta,\tag{18.163}$$

<sup>17</sup>. En réalité nous appliquons le théorème 10.1 à chacune des parties réelles et imaginaires de la fonction  $s \mapsto h'(s)\overline{h(s)}$ .

ce qu'il fallait pour l'existence.

Pour l'unicité nous supposons avoir une autre fonction,  $\alpha$  qui satisfait aux exigences. Pour tout  $t \in I$  nous avons

$$e^{i\theta(t)} = e^{i\alpha(t)}. \quad (18.164)$$

Il existe donc une fonction  $n: I \rightarrow \mathbb{N}$  telle que  $\theta(t) = \alpha(t) + 2n(t)\pi$ . Par continuité de  $\theta$  et  $\alpha$ , la fonction  $n$  doit être constante, mais vu que  $\theta(t_0) = \alpha(t_0)$  nous avons  $n = 1$ .  $\square$

### 18.5.1 Bijection continue

PROPooKSGXooOqGyZj

#### Proposition 18.59.

*L'application*

$$\begin{aligned} \gamma: [0, 2\pi[ &\rightarrow S^1 \subset \mathbb{R}^2 \\ t &\mapsto (\cos(t), \sin(t)) \end{aligned} \quad (18.165)$$

*est une bijection continue.*

*Démonstration.* En plusieurs points.

- (1) **Image** Le fait que l'image de  $\gamma$  soit dans  $S^1$  découle immédiatement du fait que  $\sin^2 + \cos^2 = 1$ .
- (2) **Continuité** La continuité découle de la continuité des composantes.
- (3) **Injectif** Soient  $x_1 < x_2$  tels que  $\sin(x_1) = \sin(x_2)$  et  $\cos(x_1) = \cos(x_2)$ . Supposons pour fixer les idées que  $\sin(x_1) > 0$  et  $\cos(x_1) > 0$  : si ce n'est pas le cas, il faut traiter séparément les 4 possibilités de combinaisons de signes.

Nous avons obligatoirement  $x_1, x_2 \in [0, \frac{\pi}{2}[$ . Vu que nous avons supposé que  $\sin(x_1) = \sin(x_2)$ , le théorème de Rolle 12.213 nous permet de considérer un élément  $c \in ]x_1, x_2[$  tel que  $\sin'(c) = 0$ , c'est-à-dire  $\cos(c) = 0$ . Cela contredirait la proposition 18.21(10) à moins que  $x_1 = x_2$ .

- (4) **Surjectif**

Soient  $x, y$  tels que  $x^2 + y^2 = 1$ . Supposons pour varier les plaisirs<sup>18</sup> que  $x < 0$  et  $y > 0$ . Puisque la fonction  $\cos$  va de 0 à  $-1$  lorsque  $x$  va de  $\pi/2$  à  $\pi$ , le théorème des valeurs intermédiaires donne  $t \in [\pi/2, \pi]$  tel que  $\cos(t) = x$ .

En reportant cette valeur de  $x$  dans l'égalité  $x^2 + y^2 = 1$ , nous trouvons  $y^2 = 1 - \cos^2(t) = \sin^2(t)$ , et donc  $\sin(t) = \pm y$ . Mais pour  $t \in [\pi/2, \pi]$  nous avons  $\sin(t) > 0$ . Par conséquent  $\sin(t) = y$  parce que nous avons également supposé  $y > 0$ .

$\square$

CORooAKMKooORqcrO

#### Corolaire 18.60.

Soient  $R > 0$  ainsi que le cercle  $S = \{x \in \mathbb{R}^2 \text{ tel que } \|x\| = R\}$  de rayon  $R$ . L'application

$$\begin{aligned} \gamma: [0, 2\pi[ &\rightarrow S \\ t &\mapsto (R \cos(t), R \sin(t)) \end{aligned} \quad (18.166)$$

*est une bijection continue.*

EXooJFDPooBZADKs

#### Exemple 18.61.

*L'application*

$$\begin{aligned} \varphi: ]0, 2\pi[ &\rightarrow S^1 \\ x &\mapsto \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix} \end{aligned} \quad (18.167)$$

est continue par la proposition 18.59. Comme  $]0, 2\pi[$  est connexe (proposition 9.54) le lemme 7.272 implique que le cercle privé d'un point est connexe.  $\triangle$

18. Mais les autres cas sont à faire, pour être complet.



La proposition suivante ne va pas jusqu'à dire qu'il y a un  $C^\infty$ -difféomorphisme entre  $[0, 2\pi[$  et  $S^1$ . La raison est que n'avons pas défini la notion de différentielle pour des applications définies sur des fermés.

Si vous voulez un  $C^\infty$  difféomorphisme, il y a deux possibilités. Soit vous acceptez d'élargir un peu le domaine et vous allez voir les coordonnées polaires du théorème 18.226, soit vous allez étudier la géométrie différentielle, et vous considérez  $S^1$  comme une variété.

PROPooXELTooYKjDav

**Proposition 18.62** ([1, 185]).

Dans cette proposition,  $S^1$  désigne l'ensemble des nombres complexes de norme 1. Trois affirmations.

ITEMooOHRHooRXvxrL

(1) *L'application*

$$\begin{aligned} f: [0, 2\pi[ &\rightarrow S^1 \\ x &\mapsto e^{ix} \end{aligned} \quad (18.168)$$

*est une bijection continue.*

ITEMooPWUWooHsQBDg

(2) *La réciproque de*

$$\begin{aligned} f: [0, 2\pi[ &\rightarrow S^1 \\ x &\mapsto e^{ix} \end{aligned} \quad (18.169)$$

*n'est pas continue en 1.*

ITEMooCVAQooLsYspI

(3) *L'application*

$$\begin{aligned} f: ]0, 2\pi[ &\rightarrow \mathbb{C} \\ x &\mapsto e^{ix} \end{aligned} \quad (18.170)$$

*est de classe  $C^\infty$ .*

*Démonstration.* En plein de points.

(1) **Pour (1)**

Nous savons que

$$\begin{aligned} \varphi: \mathbb{R}^2 &\rightarrow \mathbb{C} \\ (x, y) &\mapsto x + iy \end{aligned} \quad (18.171)$$

est une bijection isométrique. C'est pour cela que nous allons nous permettre de noter  $S^1$  le cercle unité dans  $\mathbb{R}^2$  aussi bien que l'ensemble des nombres complexes de norme 1.

Sur  $\mathbb{R}^2$  nous avons l'application

$$\begin{aligned} \gamma: [0, 2\pi[ &\rightarrow S^1 \subset \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \end{aligned} \quad (18.172)$$

qui est une bijection continue (c'est la proposition 18.59). Et enfin le lemme 18.11 nous donne  $e^{ix} = \cos(x) + i \sin(x)$ .

Avec tout ça, l'application  $\varphi^{-1} \circ f: [0, 2\pi[ \rightarrow S^1$  est une bijection continue. Et comme  $\varphi$  l'est également,  $f$  est une bijection continue.

(2) **Pour (2)** Nous vérifions que la définition 7.45(1) n'est pas satisfaite pour  $f^{-1}$ . Nous considérons le voisinage  $W = [0, \epsilon[$  de 0 dans  $[0, 2\pi[$  avec  $\epsilon < 1/2$ . Nous allons montrer que tout voisinage  $V$  de 1 dans  $S^1$  contient des points  $z_k$  tels que  $f^{-1}(z_k) \notin W$ . Tout voisinage de 1 dans  $S^1$  contient une partie de la forme  $B(1, r) \cap S^1$ . Posons  $z_k = \exp(i(2\pi - 1/k))$ . Nous avons

$$\|1 - \exp(i(2\pi - 1/k))\|^2 = \|1 - \cos(2\pi - 1/k) - i \sin(2\pi - 1/k)\|^2 \quad \text{lem 18.11} \quad (18.173a)$$

$$= |1 - \cos(2\pi - 1/k)|^2 + |\sin(2\pi - 1/k)|^2 \quad (18.173b)$$

$$\rightarrow 0. \quad (18.173c)$$

Donc pour tout  $r$ , nous avons  $z_k \in B(1, r) \cap S^1$  dès que  $k$  est assez grand. Or  $f^{-1}(z_k) = 2\pi - 1/k \notin [0, \epsilon[$  dès que  $k$  est assez grand.

- (3) **Pour (3)** Nous avons  $f(x) = \cos(x) + i \sin(x)$ . Les fonctions  $\sin$  et  $\cos$  étant de classe  $C^\infty$  par la proposition 18.1, la fonction  $f$  l'est aussi.

□

La proposition suivante donne les coordonnées polaires sur  $\mathbb{C}$ . La régularité est l'objet du théorème 18.226 (à part le fait que ce dernier parle de  $\mathbb{R}^2$  et non de  $\mathbb{C}$ ).

PROPooRFMKooURhAQJ

**Proposition 18.63** (Décomposition polaire des nombres complexes).

Pour tout nombre complexe  $z \neq 0$ , il existe un unique  $\theta \in [0, 2\pi[$  tel que

$$z = |z|e^{i\theta}. \quad (18.174)$$

*Démonstration.* Soit  $z \in \mathbb{C}$ . Nous considérons  $z' = z/|z|$  qui est de norme 1. Donc il existe un unique  $\theta \in [0, 2\pi[$  tel que  $z' = e^{i\theta}$  (proposition 18.62(1)).

Pour ce  $\theta$  nous avons  $z = |z|e^{i\theta}$ .

□

Bien entendu, le  $\theta$  est unique dans  $[0, 2\pi[$ , mais il n'est pas du tout unique dans  $\mathbb{R}$ .

LEMooOQKNooGZIJHf

**Lemme 18.64** ([1]).

Pour tout  $a, b \in \mathbb{C}$ , il existe  $t \in [0, 2\pi[$  tel que

$$|a + b|^2 = |a|^2 + |b|^2 + 2|a||b| \cos(t). \quad (18.175)$$

Plus précisément, si les formes trigonométriques<sup>19</sup> de  $a$  et  $b$  sont  $a = |a|e^{i\alpha}$  et  $b = |b|e^{i\beta}$ , alors

$$|a + b|^2 = |a|^2 + |b|^2 + 2|a||b| \cos(\alpha - \beta). \quad (18.176)$$

*Démonstration.* Nous avons le calcul suivant :

$$|a + b|^2 = (a + b)(\bar{a} + \bar{b}) \quad (18.177a)$$

$$= |a|^2 + a\bar{b} + b\bar{a} + |b|^2 \quad (18.177b)$$

$$= |a|^2 + |b|^2 + |a||b|(e^{i(\alpha-\beta)} + e^{i(\beta-\alpha)}) \quad (18.177c)$$

$$= |a|^2 + |b|^2 + 2|a||b| \cos(\alpha - \beta). \quad (18.177d)$$

□

CORooONAVooKPhQui

**Corolaire 18.65.**

La partie  $B = \{z \in \mathbb{C} \text{ tel que } |z| > R\}$  est connexe et connexe par arcs<sup>20</sup>.

*Démonstration.* Nous prouvons que  $B$  est connexe par arcs ; la connexité découlera du lemme 9.67. Nous considérons les points  $z_1 = r_1 e^{i\theta_1}$  et  $z_2 = r_2 e^{i\theta_2}$  dans  $B$ . Une telle décomposition est toujours possible par la proposition 18.63.

Supposons  $\theta_1 < \theta_2$  et  $r_1 < r_2$  ; sinon il faut faire différents cas. Nous relierons les points  $z_1$  et  $z_2$  par le chemin continue constitué des deux chemins suivants mis bout à bout. D'abord

$$\begin{aligned} \gamma: [\theta_1, \theta_2] &\rightarrow B \\ t &\mapsto r_1 e^{it}, \end{aligned} \quad (18.178)$$

et ensuite

$$\begin{aligned} \sigma: [r_1, r_2] &\rightarrow B \\ u &\mapsto u e^{i\theta_2}. \end{aligned} \quad (18.179)$$

□

19. Proposition 18.63.

20. Définition 9.65.

### 18.5.2 Inverse

SUBSECooWFNMooOuZBRN

Nous pouvons écrire un inverse de la fonction  $\varphi$  grâce à la fonction arc tangente introduite au théorème 18.39. La fonction que nous écrivons à présent est la fonction  $\arg_0$  définie par (26.137). Elle n'est pas exactement la fonction argument définie par (26.98).

Nous avons :

$$\varphi^{-1}: S^1 \rightarrow [0, 2\pi[$$

$$x + iy \mapsto \begin{cases} \arctan(y/x) & \text{si } x > 0, y \geq 0 \\ \frac{\pi}{2} & \text{si } (x, y) = (0, 1) \\ \pi - \arctan(-y/x) & \text{si } x < 0, y \geq 0 \\ \pi + \arctan(y/x) & \text{si } x < 0, y < 0 \\ 2\pi - \arctan(-y/x) & \text{si } x > 0, y < 0 \end{cases} \quad (18.180)$$

EQooSAYFooRFVSPc

Chacune des branches est continue parce que la fonction arc tangente l'est. Trois des raccords sont également continus grâce aux limites du lemme 18.40.

L'application  $\varphi^{-1}$  n'est cependant pas continue au point  $(1, 0)$ <sup>21</sup>. C'est l'objet du lemme suivant.

LEMooEQVRooMAffCw

#### Lemme 18.66.

L'application  $\varphi^{-1}: S^1 \rightarrow [0, 2\pi[$  n'est pas continue en  $(1, 0)$ . Mais elle est continue ailleurs. Autrement dit,

$$\varphi^{-1}: S^1 \setminus \{(1, 0)\} \rightarrow ]0, 2\pi[ \quad (18.181)$$

est continue.

*Démonstration.* En effet,  $\varphi^{-1}$  serait continue si l'image de tout ouvert de  $[0, 2\pi[$  par  $\varphi$  était ouverte dans  $S^1$  (topologie induite de  $\mathbb{C}$ ). Prenons un petit ouvert  $[0, \epsilon[$  (si vous êtes étonnés, c'est que vous n'avez pas bien la topologie induite en tête). Son image contient le point  $(1, 0)$ , mais aucun point  $(x, y)$  avec  $y < 0$ .

Montrons que tout voisinage de  $(1, 0)$  dans  $\mathbb{C}$  contient des points  $x + iy$  de  $S^1$  avec  $y < 0$ . Un point de  $S^1$  est de la forme  $\cos(t) + i \sin(t)$ . Nous avons :

$$|\cos(t) + i \sin(t) - 1|^2 = (\cos(t) - 1)^2 + \sin^2(t) = 2(1 - \cos(t)). \quad (18.182)$$

Soit  $\delta > 0$ , et montrons que  $B((1, 0), \delta) \cap S^1$  contient des points d'ordonnées négatives. D'abord il existe  $\epsilon > 0$  tel que pour  $t = 2\pi - \epsilon$ ,

$$2(1 - \cos(t)) < \delta. \quad (18.183)$$

Ensuite pour de tels  $t$ , nous avons  $\sin(t) < 0$ . Donc les points de  $S^1$  correspondant à  $2\pi - \epsilon$  sont dans  $S^1 \cap B((1, 0), \delta)$ .

Bref, l'image de  $[0, \epsilon[$  n'est pas un ouvert de  $S^1$ . □

### 18.5.3 Cercle trigonométrique

#### Définition 18.67 ([604]).

Le cercle trigonométrique est le cercle dans  $\mathbb{R}^2$  de rayon 1 centré en  $(0, 0)$  représenté à la figure 18.6. Nous n'hésiterons pas à parler de cercle trigonométrique dans  $\mathbb{C}$ .

Nous verrons plus tard (proposition 21.13) que la longueur de l'arc de cercle intercepté par un angle  $\theta$  est égal à  $\theta$ . Les radians sont donc l'unité d'angle la plus adaptée au calcul de longueurs sur le cercle.

<sup>21</sup>. Puisque nous avons considéré  $S^1 \subset \mathbb{C}$ , nous aurions dû noter « 1 » ce point. Mais vous vous imaginez le clash de notation avec le  $1 \in [0, 2\pi[ \subset \mathbb{R}$  ?

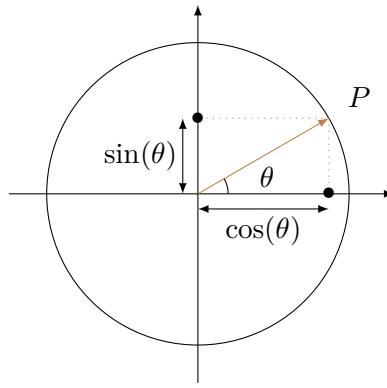


FIGURE 18.6: Le cercle trigonométrique.

LabelFigCercleTrigono

LEMooXISFooRMWUEC

**Lemme 18.68.**

Soit  $R > 0$ . Soient le cercle  $S = \{x \in \mathbb{R}^2 \text{ tel que } \|x\| = R\}$  de rayon  $R$  ainsi que  $t_0 \in \mathbb{R}$ . L'application

$$\begin{aligned} \gamma_{t_0} : [t_0, t_0 + 2\pi[ &\rightarrow S \\ x &\mapsto (R \cos(x), R \sin(x)) \end{aligned} \quad (18.184)$$

est une bijection continue.

*Démonstration.* Le corolaire 18.60 nous dit que

$$\begin{aligned} \gamma : [0, 2\pi[ &\rightarrow S \\ x &\mapsto (R \cos(x), R \sin(x)) \end{aligned} \quad (18.185)$$

est une bijection continue.

D'autre part, l'application

$$\begin{aligned} f : [t_0, t_0 + 2\pi[ &\rightarrow [0, 2\pi[ \\ x &\mapsto x - t_0 \end{aligned} \quad (18.186)$$

est une bijection continue.

Comme  $\gamma_{t_0} = \gamma \circ f$ , nous déduisons que  $\gamma_{t_0}$  est une bijection continue.  $\square$

**18.69.**

La proposition 21.13 dira que la longueur d'arc de cercle de rayon  $R$  interceptée par un angle  $\theta$  vaut  $R\theta$ .

**18.5.4 Du point de vue de la tribu, mesure et co.**

Nous avons considéré sur  $S^1$  la topologie induite de  $\mathbb{C}$ . Nous allons y mettre la tribu induite de celle de Lebesgue de  $\mathbb{C}$ . Mais nous n'allons pas y mettre la *mesure* induite de  $\mathbb{C}$ ; sinon tout serait toujours de mesure nulle.

PROPooQFYHooEajmbW

**Proposition 18.70 ([1]).**

L'application  $\varphi$  est borélienne d'inverse borélien, c'est-à-dire

$$\mathcal{Bor}(S^1) = \varphi(\mathcal{Bor}([0, 2\pi])). \quad (18.187)$$

*Démonstration.* L'inclusion  $\mathcal{Bor}(S^1) \subset \varphi(\mathcal{Bor}([0, 2\pi]))$  est la plus simple : si  $A \in \mathcal{Bor}(S^1)$ , alors  $\varphi^{-1}(A) \in \mathcal{Bor}([0, 2\pi])$  parce que  $\varphi : [0, 2\pi[ \rightarrow S^1$  est continue et donc borélienne (théorème 14.56).

Pour l'autre inclusion, il faudra procéder par étapes.

- (1) **Ouvert ne contenant pas zéro** Si  $A$  est un ouvert de  $[0, 2\pi[$  ne contenant pas 0, il est un ouvert de  $\mathbb{R}$  ou de  $]0, 2\pi[$ . Le lemme 18.66 nous indique que son image par  $\varphi$  est ouverte dans  $S^1$ . En particulier,  $\varphi(A) \in \mathcal{Bor}(S^1)$ .

- (2) **Ouvert de la forme**  $[0, \epsilon[$  Nous supposons que  $\epsilon$  est petit. Disons pour fixer les idées, plus petit que  $\pi/2$ . Nous avons :

$$\varphi([0, \epsilon[) = \varphi(]0, \epsilon[) \cup \varphi(\{0\}). \quad (18.188)$$

Le premier élément de l'union est un ouvert, et le second, un unique point. L'union est un borélien.

- (3) **Ouvert général** Si un ouvert de  $[0, 2\pi[$  ne contient pas 0, son image est ouverte. Nous nous penchons sur le cas d'un ouvert contenant 0.

Si un ouvert de  $[0, 2\pi[$  contient 0, alors il contient un ouvert de la forme  $[0, \epsilon[$ , parce qu'un ouvert contient une boule autour de chacun de ses points (théorème 7.8 couplé au fait que nous sommes dans la topologie induite de  $\mathbb{R}$ ).

Si  $A$  est un ouvert contenant zéro, alors

$$A = [0, \epsilon[ \cup (A \setminus [0, \frac{\epsilon}{2}]). \quad (18.189)$$

Nous avons déjà vu que l'image du premier élément de l'union est un borélien. Étant donné que  $A \setminus [0, \frac{\epsilon}{2}]$  est un ouvert ne contenant pas zéro, son image est un ouvert. Donc l'image de  $A$  est un borélien.

- (4) **Pause** Nous avons déjà vu que l'image par  $\varphi$  de tout ouvert de  $[0, 2\pi[$  était un borélien de  $S^1$ . Nous devons en déduire que l'image de tout borélien de  $[0, 2\pi[$  est un borélien de  $S^1$ . C'est ce que nous faisons maintenant

- (5) **Boréliens** Nous utilisons le lemme de transport 14.48 avec l'application  $\varphi^{-1}$  et l'ensemble des ouverts :

$$\varphi(\sigma(\mathcal{C})) = \sigma(\varphi(\mathcal{C})) \quad (18.190)$$

où  $\mathcal{C}$  est la tribu des ouverts dans  $[0, 2\pi[$ . L'ensemble  $\sigma(\mathcal{C})$  est par définition l'ensemble  $\mathcal{Bor}([0, 2\pi[)$ . D'autre part nous avons vu que l'image d'un ouvert est un borélien :  $\varphi(\mathcal{C}) \subset \mathcal{Bor}(S^1)$ . Nous avons donc

$$\varphi(\mathcal{Bor}([0, 2\pi[)) = \sigma(\varphi(\mathcal{C})) \subset \sigma(\mathcal{Bor}(S^1)) \subset \mathcal{Bor}(S^1). \quad (18.191)$$

La preuve est terminée. □

PROPooHMSCoorIjcJq

**Proposition 18.71** (Boréliens sur  $S^1[1]$ ).

Soit la structure usuelle d'espace mesurable  $(\mathbb{C}, \mathcal{Bor}(\mathbb{C}))$ . Nous considérons

- la tribu  $\mathcal{Bor}(\mathbb{C})_{S^1}$  induite de la tribu des boréliens de  $\mathbb{C}$  vers  $S^1$ ,
- la tribu  $\mathcal{Bor}(S^1)$  des boréliens de  $S^1$  construite à partir de la topologie induite de  $\mathbb{C}$  vers  $S^1$ .
- la bijection  $\varphi: [0, 2\pi[ \rightarrow S^1$ ,
- la mesure de Lebesgue sur  $[0, 2\pi[$  (induite de celle sur  $\mathbb{R}$ ) et sur  $\mathbb{C}$ , que nous noterons toutes deux  $\lambda$ .

Alors

ITEMooSUNEooRhAdep

- (1) Nous avons les expressions

$$\mathcal{Bor}(\mathbb{C})_{S^1} = \{A \in \mathcal{Bor}(\mathbb{C}) \text{ tel que } A \subset S^1\} \quad (18.192a)$$

$$= \{A \cap S^1 \text{ tel que } A \in \mathcal{Bor}(\mathbb{C})\} \quad (18.192b)$$

SUBEQooYZGCooDgXmft

ITEMooGYPNooRaZbNW

- (2) Nous avons

$$\mathcal{Bor}(S^1) = \mathcal{Bor}(\mathbb{C})_{S^1} = \varphi(\mathcal{Bor}([0, 2\pi[)). \quad (18.193)$$

ITEMooFUXKooFQdoaw

- (3) En définissant  $\mu: \mathcal{Bor}(S^1) \rightarrow \mathbb{R}$  par

$$\mu(A) = \frac{\lambda(\varphi^{-1}(A))}{2\pi}, \quad (18.194)$$

EQooKHZBooSrFMdo

le triplet  $(S^1, \mathcal{Bor}(S^1), \mu)$  est un espace mesuré.

ITEMooBQLRooOsquesg

(4) *L'espace mesuré  $(S^1, \mathcal{Bor}(S^1), \mu)$  est fini et*

$$\mu(S^1) = 1. \quad (18.195)$$

*Démonstration.* Point par point.

(1) **Pour (1)** C'est la proposition 14.9.

(2) **Pour (2)** La première égalité est le lemme 14.54. Le fait que  $\mathcal{Bor}(S^1) = \varphi(\mathcal{Bor}([0, 2\pi[))$  est déjà la proposition 18.70.

(3) **Pour (3)** Nous devons d'abord nous assurer que la formule ait un sens. Cela est chose aisée; si  $A \in \mathcal{Bor}(S^1)$ , le point (2) nous indique que  $\varphi^{-1}(A) \in \mathcal{Bor}([0, 2\pi[)$ . Ensuite, nous devons vérifier les deux conditions de la définition 14.18 pour avoir un espace mesuré.

En premier lieu,

$$\mu(\emptyset) = \frac{1}{2\pi} \lambda(\varphi^{-1}(\emptyset)) = \frac{1}{2\pi} \lambda(\emptyset) = 0. \quad (18.196)$$

Et en second lieu, si les  $A_i \in \mathcal{Bor}(S^1)$  sont disjoints, les  $\varphi^{-1}(A_i)$  sont également disjoints parce que  $\varphi^{-1}$  est une bijection. Donc

$$\mu\left(\bigcup_i A_i\right) = \frac{1}{2\pi} \lambda\left(\bigcup_i \varphi^{-1}(A_i)\right) \quad (18.197a)$$

$$= \frac{1}{2\pi} \sum_i \lambda(\varphi^{-1}(A_i)) \quad (18.197b)$$

$$= \sum_i \frac{\lambda(\varphi^{-1}(A_i))}{2\pi} \quad (18.197c)$$

$$= \sum_i \mu(A_i). \quad (18.197d)$$

D'accord.

(4) **Pour (4)** En ce qui concerne la mesure de  $S^1$  pour  $\mu$  nous avons simplement

$$\mu(S^1) = \frac{\lambda([0, 2\pi[)}{2\pi} = 1. \quad (18.198)$$

□

Maintenant que  $(S^1, \mathcal{Bor}(S^1), \mu)$  est un espace mesuré, nous pouvons compléter la tribu  $\mathcal{Bor}(S^1)$  pour la mesure  $\mu$ .

### Définition 18.72.

La **tribu de Lebesgue** sur  $S^1$  est la mesure complétée pour

$$(S^1, \mathcal{Bor}(S^1), \mu) \quad (18.199)$$

où  $\mu$  est la mesure définie par la proposition 18.71. Nous notons  $\mathcal{Leb}(S^1)$  la tribu et encore  $\mu$  la mesure.

PROPooDLBCooUfQZOa

### Proposition 18.73 (Lebesgue sur $S^1$ [1]).

Soit la structure d'espace mesuré complet  $(S^1, \mathcal{Leb}(S^1), \mu)$ . Nous considérons

- la tribu  $\mathcal{Leb}(\mathbb{C})_{S^1}$  induite de la tribu des boréliens de  $\mathbb{C}$  vers  $S^1$ ,
- la bijection  $\varphi: [0, 2\pi[ \rightarrow S^1$ ,

Alors

ITEMooQMHDooHEThPf

(1) La tribu  $\mathcal{Leb}(\mathbb{C})_{S^1}$  est la tribu de toutes les parties de  $S^1$ .

ITEMooNIRNooKSeYCa

(2) La tribu  $\mathcal{L}eb(S^1)$  est donnée par

$$\mathcal{L}eb(S^1) = \varphi(\mathcal{L}eb(\mathbb{R})_{[0,2\pi[}) = \varphi(\mathcal{L}eb([0, 2\pi[)) \quad (18.200)$$

où  $\mathcal{L}eb([0, 2\pi[)$  est la tribu sur  $[0, 2\pi[$  obtenue par complétion de la tribu des boréliens de la topologie induite.

ITEMooXDBTooYnaui

(3) Nous avons l'inclusion stricte

$$\mathcal{L}eb(S^1) \subsetneq \mathcal{L}eb(\mathbb{C})_{S^1}. \quad (18.201)$$

*Démonstration.* Point par point.

(1) **Pour (1)** Si  $A \subset S^1$ , alors  $A$  est une partie de  $S^1$  qui est mesurable et de mesure nulle pour  $\mathbb{C}$ . Donc  $A$  est  $\lambda$ -négligeable et par conséquent mesurable.

(2) **Pour (2)** Il s'agit de prouver que

$$\widehat{\mathcal{B}or(S^1)} = \varphi(\widehat{\mathcal{B}or([0, 2\pi[)}). \quad (18.202)$$

Ce n'est rien d'autre que la proposition 14.83. La seconde partie de l'égalité est la proposition 14.78

(3) **Pour (3)** Comme indiqué au point (1), la tribu  $\mathcal{L}eb(\mathbb{C})_{S^1}$  est la tribu de toutes les parties de  $S^1$ ; l'inclusion est donc évidente. Le point pas tout à fait évident à prouver est l'existence de parties de  $S^1$  à n'être pas dans  $\mathcal{L}eb(S^1)$ .

Soit  $V$  non mesurable dans  $[0, 2\pi[$  (prenez quelque chose comme l'ensemble de Vitali de l'exemple 14.157). Puisque, par le point (2),

$$\mathcal{L}eb(S^1) = \varphi(\mathcal{L}eb(\mathbb{R})_{[0,2\pi[}), \quad (18.203)$$

la partie  $\varphi^{-1}(V)$  ne peut pas être dans  $\mathcal{L}eb(S^1)$ .

□

Si vous en voulez plus à propos de  $S^1$  et la façon dont on passe la structure depuis  $[0, 2\pi[$ , vous pouvez lire la proposition 27.107 qui donne la structure de

$$L^2(S^1, \mathcal{L}eb(S^1), \mu) \quad (18.204)$$

qui sera, sans surprise, la même que celle de

$$L^2([0, 2\pi[, \mathcal{L}eb([0, 2\pi[), \lambda). \quad (18.205)$$

## 18.6 Exemples trigonométriques

Nous mettons ici quelques exemples concernant les fonctions trigonométriques, qui n'ont pas pu être mis dans les chapitres les plus adaptés, parce que ces derniers sont plus haut dans la table des matières.

EXooSPFDooSluUGV

### Exemple 18.74.

Prouvons que la fonction <sup>22</sup>  $f(x) = x \sin(x)$  tend vers zéro lorsque  $x$  tend vers 0. D'abord, nous coignons la fonction entre deux fonctions connues :

$$0 \leq |x \sin(x)| = |x| |\sin(x)| \leq |x|. \quad (18.206)$$

Donc  $|x \sin(x)|$  est coincé entre  $g(x) = 0$  et  $h(x) = |x|$ . Ces deux fonctions tendent vers 0 lorsque  $x \rightarrow 0$ , et donc  $f(x)$  tend vers zéro.  $\triangle$

22. La définition de la fonction sinus est 18.1.

### 18.6.1 Quelques équations trigonométriques

La proposition suivante se voit très facilement sur le cercle trigonométrique, mais il faut le démontrer.

PROP00TUUU00VrAGQ0

**Proposition 18.75** ([1]).

Si  $\theta_0 \in [0, 2\pi[$  vérifie  $\cos(\theta_0) = x_0$ , alors l'ensemble des solutions de l'équation  $\cos(\theta) = x_0$  (d'inconnue  $\theta$ ) est

$$\{\theta_0, 2\pi - \theta_0\}. \quad (18.207)$$

Cet ensemble est un singleton si et seulement si  $x_0 = \pm 1$ .

*Démonstration.* Commençons par prouver que  $\theta_0$  et  $2\pi - \theta_0$  sont des solutions. Le nombre  $\theta_0$  est solution par hypothèse. En ce qui concerne  $2\pi - \theta_0$ , il est possible d'utiliser la formule d'addition d'angle (18.22c) :

$$\cos(2\pi - \theta_0) = \cos(2\pi)\cos(\theta_0) + \sin(2\pi)\sin(\theta_0). \quad (18.208)$$

EQ00UCA000TQsUUQ

La proposition 18.21(1) nous indique que  $\cos(2\pi) = 1$  et  $\sin(2\pi) = 0$ . Donc l'égalité (18.208) se réduit à  $\cos(2\pi - \theta_0) = \cos(\theta_0)$ .

Le lemme 18.4 dit que si  $\cos(\theta) = x_0$ , alors

$$\sin(\theta) = \pm\sqrt{1 - x_0^2}. \quad (18.209)$$

Nous avons donc soit

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} x_0 \\ \sqrt{1 - x_0^2} \end{pmatrix}, \quad (18.210)$$

soit

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} x_0 \\ -\sqrt{1 - x_0^2} \end{pmatrix}, \quad (18.211)$$

Comme  $\theta \mapsto (\cos(\theta), \sin(\theta))$  est une bijection avec  $S^1$  (proposition 18.59), chacune de ces deux possibilités possède une unique solution. L'ensemble des solutions de  $\cos(\theta) = x_0$  possède donc au maximum deux éléments.

L'ensemble des solutions possède exactement une solution lorsque les points  $(x_0, \sqrt{1 - x_0^2})$  et  $(x_0, -\sqrt{1 - x_0^2})$  sont identiques. C'est le cas, si et seulement si  $\sqrt{1 - x_0^2} = 0$ , c'est-à-dire, si et seulement si  $x_0 = \pm 1$ .  $\square$

### 18.6.2 Développements en série

PROP00NPYX00TuwAHP

**Proposition 18.76** (Taylor pour cosinus).

Pour tout  $n \in \mathbb{N}$ , il existe une fonction  $\alpha: \mathbb{R} \rightarrow \mathbb{R}$  telle que  $\lim_{t \rightarrow 0} \alpha(t) = 0$  et

$$\cos(x) = \sum_{k=0}^n \frac{(-1)^k}{(2k)!} x^{2k} + \alpha(x)x^{2n}. \quad (18.212)$$

EQ00GQOI00IkwbJIV

En ce qui concerne le sinus, pour tout  $n$  nous avons une fonction  $\alpha: \mathbb{R} \rightarrow \mathbb{R}$  telle que  $\lim_{t \rightarrow 0} \alpha(t) = 0$  et

$$\sin(x) = \sum_{k=0}^n \frac{(-1)^k}{(2k+1)!} x^{2k+1} + x^{2n+2}\alpha(x). \quad (18.213)$$

EQ00KYJA00RebHgc

*Démonstration.* Il s'agit d'utiliser la proposition 15.57, en faisant attention à l'ordre. Le fait est que dans (18.212), nous avons écrit le polynôme de degré  $2n+1$  (et non seulement  $2n$ ), en sachant que le terme d'ordre  $2n+1$  est nul.

C'est pour cela que nous avons pu écrire  $\alpha(x)x^{2n+1}$  au lieu de  $\alpha(x)x^{2n}$  qui aurait été attendu.

Même raisonnement pour le développement du sinus.  $\square$

**Remarque 18.77.**

Quelques remarques concernant l'ordre du polynôme.



- (1) Notons que nous aurions aussi pu écrire le reste sous la forme  $\alpha(x)x^{2n}$ , mais ça aurait été avec une autre fonction  $\alpha$  : celle correspondant au développement à l'ordre  $2n$  au lieu de  $2n+1$ .
- (2) Les développements de sinus et de cosinus ont un terme sur deux qui est nul. C'est pour cela qu'en ayant un polynôme de degré  $2p$ , nous avons le développement d'ordre  $2p+1$ .
- (3) Nous aurions pu utiliser les dérivées données dans la proposition 18.3 et les valeurs spéciales (18.6).

**Corolaire 18.78.**

Il existe une fonction  $\alpha: \mathbb{R} \rightarrow \mathbb{R}$  telle que  $\lim_{t \rightarrow 0} \alpha(t)/t = 0$  et

$$\sin(x) = x + \alpha(x). \quad (18.214)$$

Nous avons la limite

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = 1. \quad (18.215)$$

*Démonstration.* Il s'agit de prendre la formule (18.213) avec  $n = 0$ . Cela donne tout de suite (18.214). Pour la limite, on divise par  $x$ , ce qui donne (pour tout  $x \neq 0$ )

$$\frac{\sin(x)}{x} = 1 + \frac{\alpha(x)}{x}. \quad (18.216)$$

Et justement la fonction  $\alpha$  a la propriété que  $\lim_{x \rightarrow 0} \alpha(x)/x = 0$ .  $\square$

**Exemple 18.79.**

Cherchons le développement limité à l'ordre 5 de  $\tan(x) = \frac{\sin(x)}{\cos(x)}$ . Nous utilisons les développements de la proposition 18.76 :

$$\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} + x^5 \alpha_1(x) \quad (18.217a)$$

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} + x^5 \alpha_2(x). \quad (18.217b)$$

Nous calculons alors la division des deux polynômes, en classant les puissances dans l'ordre croissant (c'est le sens inverse de ce qui est fait pour la divisions euclidienne!) :

$$\begin{array}{r|l} \begin{array}{r} x \quad - \quad \frac{1}{6}x^3 \quad + \quad \frac{1}{120}x^5 \\ x \quad - \quad \frac{1}{2}x^3 \quad + \quad \frac{1}{24}x^5 \\ \hline \frac{1}{3}x^3 \quad - \quad \frac{1}{30}x^5 \\ - \left( \frac{1}{3}x^3 \quad - \quad \frac{1}{6}x^5 \quad + \quad \frac{1}{72}x^7 \right) \\ \hline \frac{2}{15}x^5 \quad - \quad \frac{1}{72}x^7 \\ - \left( \frac{2}{15}x^5 \quad - \quad \frac{1}{15}x^7 \quad + \quad \frac{1}{180}x^9 \right) \\ \hline \frac{29}{360}x^7 \quad - \quad \frac{1}{180}x^9 \end{array} & \begin{array}{l} 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 \\ \hline x + \frac{1}{3}x^3 + \frac{2}{15}x^5 \end{array} \end{array}$$

Nous avons continué la division jusqu'à obtenir un reste de degré plus grand que 5. Le développement à l'ordre 5 de la fonction tangente autour de zéro est alors (proposition 12.522)

$$\tan(x) = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + x^5 \alpha(x). \quad (18.218)$$

Notons que, puisque le reste ne nous intéressait pas vraiment, nous aurions pu ne pas calculer les coefficients des termes en  $x^7$  et  $x^9$ . La dernière soustraction était également inutile.  $\triangle$

## 18.7 Isométries de l'espace euclidien

Nous considérons l'espace affine euclidien  $A = \mathcal{E}_n(\mathbb{R})$  modelé sur  $\mathbb{R}^n$  avec sa métrique usuelle. Un premier grand résultat sera le théorème 8.154 qui dira que les isométries de cet espace sont des applications linéaires.

### 18.7.1 Structure du groupe $\text{Isom}(\mathbb{R}^n)$

Si vous ne voulez pas savoir ce qu'est un produit semi-direct de groupes, vous pouvez lire seulement le point (1) du théorème suivant, et passer directement à la remarque 18.81.

THOooQJSRooMrqQct

#### Théorème 18.80.

Un peu de structure sur  $\text{Isom}(\mathbb{R}^n)$ .

ITEMooLLUIooIGsknv

(1) L'application

$$\begin{aligned}\psi: T(n) \times O(n) &\rightarrow \text{Isom}(\mathbb{R}^n) \\ (v, \Lambda) &\mapsto \tau_v \circ \Lambda\end{aligned}\tag{18.219}$$

est une bijection. Ici,  $T(n)$  est le groupe des translations de  $\mathbb{R}^n$ .

(2) Un couple  $(v, \Lambda) \in T(n) \times SO(n)$  agit sur  $x \in \mathbb{R}^n$  par

$$(v, \Lambda)x = \Lambda x + v\tag{18.220}$$

au sens où  $\psi(v, \Lambda)x = \Lambda x + v$ .

ITEMooEWSIooNKzRxkB

(3) En tant que groupes,

$$\text{Isom}(\mathbb{R}^n) \simeq T(n) \times_{\rho} O(n)\tag{18.221}$$

où  $\rho$  représente l'action adjointe de  $O(n)$  sur  $T(n)$  et  $\times_{\rho}$  dénote le produit semi-direct de la définition 2.48.

ITEMooSKUPooBDvNWx

(4) Une isométrie de  $\mathbb{R}^n$  est une application affine<sup>23</sup>.

ITEMooQLNPooSyHaps

(5) La partie linéaire<sup>24</sup> d'une isométrie  $f$  est  $\psi^{-1}(f)_2$ .

Démonstration. Point par point.

(1) Prouvons que l'application proposée est injective et surjective. Notons aussi que ce point ne parle pas de structure de groupe, mais seulement d'une bijection en tant qu'ensembles.

(1a) **Injection** Si  $\psi(v, \Lambda) = \psi(w, \Lambda')$  alors en appliquant sur  $x = 0$  nous avons tout de suite  $v = w$ . Et ensuite  $\Lambda = \Lambda'$  est immédiat.

(1b) **Surjection** Une isométrie  $g \in \text{Isom}(\mathbb{R}^n)$  est une application  $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$  telle que  $d(x, y) = d(g(x), g(y))$ . Dans le cas de  $\mathbb{R}^n$  cela se traduit par

$$\|x - y\| = \|g(x) - g(y)\|,\tag{18.222}$$

Comme  $x \mapsto \|x\|^2$  est une forme quadratique, elle tombe sous le coup du théorème 8.154, ce qui nous permet de dire que  $g$  est affine. Or par définition une application est affine lorsqu'elle est la composée d'une translation et d'une application linéaire.

Donc  $g = \tau_v \circ \Lambda$  pour une certaine application linéaire isométrique  $\Lambda: \mathbb{R}^n \rightarrow \mathbb{R}^n$ . L'application  $\Lambda$  est donc dans  $O(n)$  par la proposition 8.45(3).

(2) C'est seulement le fait que  $(\tau_v \circ \Lambda)x = \tau_v(\Lambda x) = \Lambda(x) + v$ .

(3) Nous allons étudier l'application

$$\psi: T(n) \times_{\rho} O(n) \rightarrow \text{Isom}(\mathbb{R}^n).\tag{18.223}$$

(3a) **Le produit semi-direct est bien défini** Il faut montrer que

$$\begin{aligned}\rho: O(n) &\rightarrow \text{Aut}(T(n)) \\ \Lambda &\mapsto \mathbf{Ad}(\Lambda)\end{aligned}\tag{18.224}$$

est correcte.

23. Définition 11.13.

24. Définition 11.14.

D'abord pour  $\Lambda \in O(n)$ , nous avons bien  $\rho_\Lambda(\tau_v) \in T(n)$  parce qu'en appliquant à  $x \in \mathbb{R}^n$ ,

$$(\Lambda\tau_v\Lambda^{-1})(x) = \Lambda(\tau_v(\Lambda^{-1}x)) = \Lambda(\Lambda^{-1}x + v) = x + \Lambda(v) = \tau_{\Lambda(v)}(x). \quad (18.225)$$

Donc  $\rho_\Lambda(\tau_v) = \tau_{\Lambda(v)}$ .

De plus,  $\rho_\Lambda \in \text{Aut}(T(n))$  parce que

$$\rho_\Lambda(\tau_v \circ \tau_w) = \rho_\Lambda(\tau_v) \circ \rho_\Lambda(\tau_w), \quad (18.226)$$

comme on peut aisément vérifier que les deux membres sont égaux à  $\tau_{\Lambda(v+w)}$ .

(3b)  $\psi$  est une bijection Cela est déjà vérifié.

(3c)  $\psi$  est un morphisme Nous avons d'une part

$$\psi((v, g)(w, h)) = \psi(v\rho_g(w), gh) = \tau_v \circ g \circ \tau_w \circ g^{-1} \circ g \circ h = \tau_v \circ g \circ \tau_w \circ h. \quad (18.227)$$

Et d'autre part,

$$\psi(v, g) \circ \psi(w, h) = \tau_v \circ g \circ \tau_w \circ h, \quad (18.228)$$

ce qui est la même chose.

(4) Si  $f$  est une isométrie de  $\mathbb{R}^n$ , et si  $\psi(f) = (v, \Lambda)$ , nous avons

$$f(a + x) = \tau_v(\Lambda(a) + \Lambda(x)) = \Lambda(a) + \Lambda(x) + v = f(a) + \Lambda(x). \quad (18.229)$$

Donc  $f$  est affine et l'application qui serait notée  $u_M$  dans (11.42) est  $u_M = \Lambda$  pour tout  $M$ .

(5) Si  $f(x) = \psi(v, \Lambda)x = \Lambda x + v$ , nous avons

$$f(a + x) = \psi(v, \Lambda)(a + x) = f(a) + \Lambda(x), \quad (18.230)$$

et donc  $\Lambda$  est bien la partie linéaire de  $f$ .

□

REMooLUEZooIwvTqu

### Remarque 18.81.

Notons au passage la loi de groupe sur les couples qui est donnée, pour tout  $v, v' \in \mathbb{R}^n$ ,  $\Lambda, \Lambda' \in \text{SO}(n)$ , par

$$(v, \Lambda) \cdot (v', \Lambda') = (\Lambda v' + v, \Lambda \Lambda') \quad (18.231) \quad \text{EgDiHcut}$$

comme le montre le calcul suivant :

$$(v, \Lambda) \cdot (v', \Lambda')x = (v, \Lambda)(\Lambda'x + v') \quad (18.232a)$$

$$= \Lambda \Lambda'x + \Lambda v' + v \quad (18.232b)$$

$$= (\Lambda v' + v, \Lambda \Lambda')x. \quad (18.232c)$$

PROPooDHYWooXxEXvl

### Proposition 18.82 ([605]).

Soient  $n \geq 1$  et  $R$  un élément de  $O(n)$  de déterminant  $-1$  tels que  $R^2 = \text{Id}$ . En posant  $C_2 = \{\text{Id}, R\}$  nous avons

$$O(n) = \text{SO}(n) \times_\rho C_2 \quad (18.233)$$

*Démonstration.* Notons qu'un élément  $R$  comme décrit dans l'énoncé existe. Par exemple il y a l'application  $(x_1, \dots, x_n) \mapsto (-x_1, x_2, \dots, x_n)$ .

Cela étant dit, nous allons montrer que

$$\begin{aligned} \psi: \text{SO}(n) \times C_2 &\rightarrow O(n) \\ (A, h) &\mapsto Ah. \end{aligned} \quad (18.234)$$

est un isomorphisme.

- (1) **Injectif** Soient  $A, B \in \text{SO}(n)$  et  $h, k \in C_2$  tels que  $\psi(A, h) = \psi(B, k)$ , c'est-à-dire tels que  $Ah = Bk$ . Puisque  $\det(A) = \det(B) = 1$ , nous avons  $\det(h) = \det(k)$ . Mais comme  $C_2$  contient un élément de déterminant 1 et un élément de déterminant  $-1$ , nous avons  $h = k$ . De là  $A = B$ .
- (2) **Surjectif** Soit  $X \in \text{O}(n)$ . Si  $\det(X) = 1$  alors  $X \in \text{SO}(n)$  et  $X = \psi(X, \mathbb{1})$ . Si par contre  $\det(X) = -1$ , alors  $XR \in \text{SO}(n)$  parce que  $\det(XR) = 1$ , et nous avons

$$\psi(XR, R) = XR^2 = X. \quad (18.235)$$

- (3) **Morphisme** Nous avons

$$\psi((A, h)(B, k)) = \psi(A\rho_h(B), hk) = A(hBh^{-1})hk = AhBk, \quad (18.236)$$

tandis que

$$\psi(A, h)\psi(B, k) = AhBk, \quad (18.237)$$

qui est la même chose.

□

## 18.8 Isométries dans $\mathbb{R}^n$

### Définition 18.83.

Un **hyperplan** de  $\mathbb{R}^n$  est un sous-espace affine de dimension  $n - 1$ .

LEM00PXIO00SIMIHY

### Lemme-Définition 18.84.

Si un hyperplan  $H$  de  $\mathbb{R}^n$  est donné, et si  $x \in \mathbb{R}^n$ , il existe un unique point  $y \in \mathbb{R}^n$  tel que

$$(1) \quad x - y \perp H,$$

(2) Le segment  $[x, y]$  coupe  $H$  en son milieu.

La **réflexion**  $\sigma_H$  est l'application  $\sigma_H: \mathbb{R}^n \rightarrow \mathbb{R}^n$  qui à  $x$  fait correspondre ce  $y$ .

*Démonstration.* Il faut vérifier que les conditions données définissent effectivement un unique point de  $\mathbb{R}^n$ . Soit  $H_0$  le sous-espace vectoriel parallèle à  $H$  et une base orthonormée  $\{e_1, \dots, e_{n-1}\}$  de  $H_0$ . Nous complétons cette partie en une base orthonormée de  $\mathbb{R}^n$  avec un vecteur  $e_n$ . Si  $H = H_0 + v$ , quitte à décomposer  $v$  en une partie parallèle et une partie perpendiculaire à  $H$ , nous avons

$$H = H_0 + \lambda e_n \quad (18.238)$$

pour un certain  $\lambda$ .

Une droite passant par  $x$  et perpendiculaire à  $H$  est de la forme  $t \mapsto x + te_n$ . Si  $x = \sum_{i=1}^n x_i e_i$  alors l'unique point de cette droite à être dans  $H$  est le point tel que  $x_n e_n + te_n = \lambda e_n$ , c'est-à-dire  $t = \lambda - x_n$ . L'unique point  $y$  sur cette droite à être tel que  $[x, y]$  coupe  $H$  en son milieu est celui qui correspond à  $t = 2(\lambda - x_n)$ . □

### Lemme 18.85.

Soit un hyperplan vectoriel  $H_0$  de  $\mathbb{R}^n$ . Nous considérons une base  $\{e_1, \dots, e_n\}$  de  $\mathbb{R}^n$  telle que  $\{e_1, \dots, e_{n-1}\}$  soit une base de  $H_0$  et  $e_n \perp H_0$ .

Soit l'hyperplan affine  $H = H_0 + \lambda e_n$ .

La réflexion de plan  $H$  est donnée par

$$\begin{aligned} \sigma: \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\mapsto x - 2(x_n - \lambda)e_n. \end{aligned} \quad (18.239)$$

EQ00RTWLe0LPsUpY

*Démonstration.* En posant  $y = x - 2(x_n - \lambda)e_n$ , nous avons

$$x - y = 2(x_n - \lambda)e_n \perp H, \quad (18.240)$$

et

$$\frac{x + y}{2} = \frac{2x - 2(x_n - \lambda)e_n}{2} = x - (x_n - \lambda)e_n = \sum_{i=1}^{n-1} x_i e_i + \lambda e_n \in H_0 + \lambda e_n = H. \quad (18.241)$$

Les deux conditions de la définition 18.84 sont donc vérifiées.  $\square$

LEMooWYVRooQmWqvM

**Lemme 18.86.**

Si  $H_0$  est un plan vectoriel, si  $\{e_1, \dots, e_{n-1}\}$  est une base de  $H_0$ , si  $e_n \perp H_0$  et si  $H = H_0 + \lambda e_n$ , alors

$$\sigma_H = \sigma_{H_0} + 2\lambda e_n. \quad (18.242)$$

*Démonstration.* En posant  $\lambda = 0$  dans la formule (18.239), nous avons  $\sigma_{H_0}(x) = x - 2x_n e_n$ . Nous avons donc

$$\sigma_H(x) = x + 2(\lambda - x_n)e_n = x - 2x_n e_n + 2\lambda e_n = \sigma_{H_0}(x) + 2\lambda e_n. \quad (18.243)$$

$\square$

Le lemme suivant est une généralisation du fait que tous les points de la médiatrice d'un segment sont à égale distance des deux extrémités du segment (très utile lorsqu'on étudie les triangles isocèles).

LEMooDPLYooJKZxiM

**Lemme 18.87** ([605]).

Soient deux points distincts  $x_0, y_0 \in \mathbb{R}^n$  l'ensemble  $H \subset \mathbb{R}^n$  donné par

$$H = \{x \in \mathbb{R}^n \text{ tel que } d(x, x_0) = d(x, y_0)\}. \quad (18.244)$$

Alors

- (1)  $H$  est hyperplan.
- (2)  $H \perp y_0 - x_0$
- (3)  $H$  contient le milieu du segment  $[x_0, y_0]$ .

*Démonstration.* Nous savons que

$$d(x, x_0)^2 = \langle x - x_0, x - x_0 \rangle = \|x\|^2 + \|x_0\|^2 - 2\langle x, x_0 \rangle, \quad (18.245)$$

ou encore

$$\|x_0\|^2 - \|y_0\|^2 = 2\langle x, x_0 - y_0 \rangle. \quad (18.246)$$

En posant  $v = y_0 - x_0$  et en considérant la forme linéaire

$$\begin{aligned} \beta: \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto \langle x, v \rangle, \end{aligned} \quad (18.247)$$

Nous avons  $x \in H$  si et seulement si  $\beta(x) = \frac{1}{2}(\|y_0\|^2 - \|x_0\|^2) = \lambda$ . En d'autres termes,  $H = \beta^{-1}(\lambda)$ . Par la proposition 11.28 la partie  $H$  est un sous-espace affine. C'est même un translaté de  $\ker(\beta)$ , et comme  $\ker(\beta)$  est l'espace vectoriel des vecteurs perpendiculaires à  $v$ , nous avons  $\dim(H) = \dim(\ker(\beta)) = n - 1$ .

Le fait que  $H$  contienne le milieu du segment  $[x_0, y_0]$  est par définition.  $\square$

Pour le lemme suivant, et pour que la récurrence se passe bien nous disons que l'ensemble vide est un espace vectoriel de dimension  $-1$ .

LEMooJCdRooGAmlwp

**Lemme 18.88** (Cartan-Dieudonné[606, 607]).

Soit un espace euclidien  $E$  de dimension  $n$ .

ITEMooFYEDooIJZBjP

(1) Si  $f$  est une isométrie de  $E$  satisfaisant

$$\dim(\text{Fix}(f)) = n - k \quad (18.248)$$

alors  $f$  peut être écrit comme composition de  $k$  réflexions hyperplanes.

ITEMooJTZVooWvyfDD

(2) Une isométrie de  $E$  peut être écrite sous la forme de  $\text{rk}(f - \text{Id})$  réflexions, mais pas moins.

ITEMooUoCZWooSbyPwt

(3) Toute isométrie de  $E$  peut être écrite comme composition de  $n + 1$  réflexions.

*Démonstration.* Les deux parties importantes à démontrer sont les points (1) et la partie « pas moins » de (2). Le reste proviendra de reformulations.

(1) **Pour (1)** Nous faisons une récurrence sur  $k \geq 0$ .

Pour l'initialisation, si  $k = 0$  alors  $\dim(\text{Fix}(f)) = n$ , c'est-à-dire que  $f$  fixe tout  $\mathbb{R}^n$ , autant dire que  $f$  est l'identité, une composition de zéro réflexions.

Pour la récurrence, nous supposons que le lemme est démontré jusqu'à  $k \geq 0$ . Soit donc  $f \in \text{Isom}(\mathbb{R}^n)$  tel que

$$\dim(\text{Fix}(f)) = n - (k + 1). \quad (18.249)$$

Puisque  $k \geq 0$ , la dimension de  $\text{Fix}(f)$  est strictement plus petite que  $n$ , donc il existe un  $x_0 \in \mathbb{R}^n$  tel que  $f(x_0) \neq x_0$ . Nous posons

$$H = \{x \in E \text{ tel que } d(x, x_0) = d(x, f(x_0))\}. \quad (18.250)$$

Par le lemme 18.87, ce  $H$  est l'hyperplan orthogonal à  $v = f(x_0) - x_0$  et passant par le milieu du segment  $[x_0, f(x_0)]$ .

Nous posons  $g = \sigma_H \circ f$ . Comme  $g(x_0) = \sigma_H(f(x_0)) = x_0$ , ce  $x_0$  est un point fixe de  $g$ . Le fait que  $\sigma_H(f(x_0)) = x_0$  est vraiment la définition de l'hyperplan  $H$ .

Nous avons donc

$$x_0 \in \text{Fix}(g) \setminus \text{Fix}(f). \quad (18.251)$$

Mais nous prouvons de plus que  $\text{Fix}(f) \subset \text{Fix}(g)$ . En effet si  $y \in \text{Fix}(f)$  alors  $y \in H$  parce que

$$d(y, x_0) = d(f(y), f(x_0)) = d(y, f(x_0)). \quad (18.252)$$

Puisque  $y \in H$  nous avons  $y \in \text{Fix}(g)$  parce que

$$g(y) = \sigma_H(f(y)) = \sigma_H(y) = y. \quad (18.253)$$

Tout cela pour dire que l'ensemble  $\text{Fix}(g)$  est *strictement* plus grand que  $\text{Fix}(f)$ . Et comme ce sont des espaces affines, nous pouvons parler de dimension :

$$\dim(\text{Fix}(g)) > \dim(\text{Fix}(f)). \quad (18.254)$$

Par hypothèse de récurrence, l'application  $g$  peut être écrite comme composition de  $k$  réflexions. Donc l'application

$$f = \sigma_H \circ g \quad (18.255)$$

est une composition de  $k + 1$  réflexions.

(2) **Pour (2), existence**

Ce point est une reformulation du point (1). Le fait est que  $\text{Fix}(f) = \ker(f - \text{Id})$  parce que  $x \in \text{Fix}(f)$  si et seulement si  $f(x) = x$  si et seulement si  $(f - \text{Id})x = 0$ . Nous utilisons le théorème du rang 4.49 à l'endomorphisme  $f - \text{Id}$  :

$$\dim(\text{Fix}(f)) = \dim(\ker(f - \text{Id})) = \dim(E) - \text{rk}(f - \text{Id}). \quad (18.256)$$

En remplaçant par les valeurs :

$$n - k = n - \text{rk}(f - \text{Id}). \quad (18.257)$$

Or le point (1) donnait  $f$  comme composée de  $n - k$  réflexions. Donc  $f$  est composée de  $\text{rk}(f - \text{Id})$  réflexions.

(3) **Pour (2), « pas moins »**

Supposons que  $f = \sigma_1 \circ \dots \circ \sigma_r$  où  $\sigma_i$  est la réflexion de l'hyperplan  $H_i$ . Nous devons prouver que  $r \geq \text{rk}(f - \text{Id})$ . Nous avons

$$\bigcap_{i=1}^r H_i \subset \ker(f - \text{Id}). \quad (18.258)$$

D'autre part, la proposition 8.308 nous donne  $\dim \bigcap_i H_i \geq n - r$ . Donc

$$n - r \leq \dim \left( \bigcap_{i=1}^r H_i \right) \leq \dim (\ker(f - \text{Id})) = n - \text{rk}(f - \text{Id}). \quad (18.259)$$

Donc  $n - r \leq n - \text{rk}(f - \text{Id})$  ou encore

$$r \geq \text{rk}(f - \text{Id}). \quad (18.260)$$

(4) **Pour (3)** Le rang de  $f - \text{Id}$  vaut au maximum  $n$ . Donc  $f$  peut être écrite comme composition de  $n + 1$  réflexions par le point (2). □

PROPooUSKEooUbNVfs

**Proposition 18.89.**

Un élément de  $\text{SO}(3)$  qui fixe deux vecteurs linéairement indépendants est l'identité.

*Démonstration.* Soit un élément  $A \in \text{SO}(3)$  et deux vecteurs linéairement indépendants  $v_1, v_2 \in \mathbb{R}^3$  tels que  $Av_1 = v_1$  et  $Av_2 = v_2$ . Puisque  $v_1$  et  $v_2$  sont linéairement indépendants, le théorème de la base incomplète 4.14 nous permet de considérer  $v_3 \in \mathbb{R}^3$  tel que  $\{v_1, v_2, v_3\}$  soit une base. Dans cette base, la matrice de  $A$  est de la forme

$$A = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}. \quad (18.261)$$

Le déterminant de cette matrice est  $c$ . Or  $\det(A) = 1$  parce qu'elle est dans  $\text{SO}(3)$ . Donc  $c = 1$ . Le fait que  $A$  soit orthogonale implique que la troisième colonne doit être un vecteur de norme 1. Donc  $a = b = 0$ .

Donc  $A = \text{Id}$ . □

CORooJCURooSRzSFb

**Corolaire 18.90.**

Tout élément de  $\text{SO}(3)$  peut être écrit comme composée de deux réflexions.

*Démonstration.* Un élément de  $\text{SO}(3)$  est une isométrie de  $\mathbb{R}^3$  parce que si  $A \in \text{SO}(3)$  alors <sup>25</sup>

$$\langle Ax, Ay \rangle = \langle A^* Ax, y \rangle = \langle x, y \rangle. \quad (18.262)$$

Donc si le rang de  $A$  est  $k$ , alors  $A$  est la composée de  $3 - k$  réflexions par le lemme 18.88.

Si  $A = \text{Id}$ , c'est bon parce que l'identité est la composée de deux réflexions égales. Nous supposons que  $A$  n'est pas l'identité.

Comme discuté dans l'exemple 12.94, l'opérateur  $A$  possède trois valeurs propres dans  $\mathbb{C}$  dont une réelle, et deux complexes conjuguées. Nous les notons  $\lambda \in \mathbb{R}$  et  $\alpha, \bar{\alpha} \in \mathbb{C}$ . Le déterminant de  $A$ , qui vaut 1, est le produit de ces trois valeurs propres, c'est-à-dire  $\lambda|\alpha|^2$ . En particulier  $\lambda > 0$ .

Si  $v$  est un vecteur propre correspondant à la valeur propre  $\lambda$ , nous avons  $\|v\| = \|Av\| = |\lambda|\|v\|$  parce que  $A$  est une isométrie. Donc  $\lambda = \pm 1$ .

Au final,  $\lambda = 1$ . Cela signifie que  $A$  laisse au moins un vecteur invariant. Vu que  $A$  n'est pas l'identité, la proposition 18.89 nous indique qu'il n'y a pas d'autres vecteurs de  $\mathbb{R}^3$  à être fixé par  $A$ . Donc  $\dim(\text{Fix}(A)) = 1$  et le lemme 18.88(1) s'écrit avec  $n = 3$ ,  $k = 2$  et implique que  $A$  est la composée de deux réflexions. □

25. Opérateur orthogonal, définition 8.42.

LEMooMCVKooKzmlAg

**Lemme 18.91.**

Soit un hyperplan  $H$  et un vecteur  $v$  de  $\mathbb{R}^n$ . Nous avons

$$\tau_v \circ \sigma_H \circ \tau_v^{-1} = \sigma_{\tau_v(H)}. \quad (18.263)$$

*Démonstration.* Pour ce faire, nous considérons une base adaptée. Les vecteurs  $\{e_1, \dots, e_{n-1}\}$  forment une base orthonormée de  $H_0$  et  $e_n$  complète en une base orthonormée de  $\mathbb{R}^n$ . Soit  $H_0$  l'hyperplan parallèle à  $H$  et passant par l'origine ; nous avons, pour un certain  $\lambda \in \mathbb{R}$ ,

$$H = H_0 + \lambda e_n \quad (18.264)$$

D'un autre côté, le vecteur  $v$  peut être décomposé en  $v = v_1 + v_2$  où  $v_1 \perp H$  et  $v_2 \parallel H$ . Alors

$$\tau_v(H) = H + v = H + v_2 = H_0 + \lambda e_n + v_2. \quad (18.265)$$

Nous pouvons maintenant utiliser le lemme 18.86 pour exprimer la transformation  $\sigma_{\tau_v(H)}$  :

$$\sigma_{\tau_v(H)}(x) = \sigma_{H_0}(x) + 2\lambda e_n + 2v_2 \quad (18.266) \quad \text{EQooNYKFooXprXav}$$

Mais d'autre part,

$$(\tau_v \circ \sigma_H \circ \tau_v^{-1})(x) = v + \sigma_H(x - v) = v + \sigma_{H_0}(x - v) + 2\lambda e_n. \quad (18.267)$$

Vue la décomposition de  $v = v_1 + v_2$  nous avons  $\sigma_{H_0}(v) = -v_1 + v_2$  et donc

$$(\tau_v \circ \sigma_H \circ \tau_v^{-1})(x) = v + \sigma_{H_0}(x) + v_1 - v_2 + 2\lambda e_n = \sigma_{H_0}(x) + 2v_1 + 2\lambda e_n. \quad (18.268) \quad \text{EQooGOHEooALPBBB}$$

Les expressions (18.266) et (18.268) coïncident, d'où l'égalité recherchée.  $\square$

PROPooLYCUooRQgGtF

**Proposition 18.92.**

Soit un hyperplan  $H$  de  $\mathbb{R}^n$  passant par l'origine. Il existe une base orthonormée  $\{e_1, \dots, e_n\}$  de  $\mathbb{R}^n$  telle que  $\{e_1, \dots, e_{n-1}\}$  soit une base orthonormée de  $H$  et  $e_n \perp H$ .

**18.8.1 Préserver l'orientation**

Le fait qu'une isométrie puisse être décomposé en réflexions est le lemme 18.88.

DEFooUZFHooXVVLBL

**Proposition-Définition 18.93** (Isométrie positive et négative[605]).

Nous utilisons la bijection  $\psi: T(n) \times O(n) \rightarrow \text{Isom}(\mathbb{R}^n)$  décrite par le théorème 18.80, et nous définissons

$$\begin{aligned} \epsilon: \text{Isom}(\mathbb{R}^n) &\rightarrow \{\pm 1\} \\ f &\mapsto \det(\psi^{-1}(f)_2). \end{aligned} \quad (18.269)$$

Cette application  $\epsilon$  est un morphisme de groupes.

Nous disons que  $f \in \text{Isom}(\mathbb{R}^n)$  est **positive** ou **préserve l'orientation** si  $\epsilon(f) = 1$ , et nous notons

$$\text{Isom}^+(\mathbb{R}^n) = \epsilon^{-1}(1). \quad (18.270)$$

LEMooVRELooESIWQi

**Lemme 18.94.**

La partie  $\text{Isom}^+(\mathbb{R}^n)$  des isométries positives est un sous-groupe de  $\text{Isom}(\mathbb{R}^n)$ .

LEMooJABDooOKHwWv

**Lemme 18.95** ([605]).

Si  $H$  est un hyperplan et si  $\sigma_H$  est la réflexion de cet hyperplan, alors  $\epsilon(\sigma_H) = -1$ .

*Démonstration.* Nous commençons par supposer que  $H$  passe par l'origine, de telle sorte que  $\sigma_H$  soit linéaire, et que  $\epsilon(\sigma_H) = \det(\sigma_H)$ .



Nous choisissons une base orthonormée de  $\mathbb{R}^n$  comme dans la proposition 18.92 :  $\{e_1, \dots, e_{n-1}\}$  est une base de  $H$  et  $e_n \perp H$ . Nous avons alors

$$\sigma_H(e_i) = \begin{cases} e_i & \text{si } i = 1, \dots, n-1 \\ -e_n & \text{si } i = n. \end{cases} \quad (18.271)$$

Calculons ce déterminant à l'ancienne, en utilisant les définitions 8.15 et 8.10. Nous posons  $\mu_i = 1$  pour  $i = 1, \dots, n-1$  et  $\mu_n = -1$ , de telle sorte à avoir  $\sigma_H(e_i) = \mu_i e_i$ . Nous posons aussi  $B = (e_1, \dots, e_n)$ . Ensuite c'est parti pour le calcul :

$$\det_B(f(B)) = \sum_{s \in S_n} \epsilon(s) \prod_{i=1}^n e_{s(i)}^*(f(e_i)) \quad (18.272a)$$

$$= \sum_{s \in S_n} \epsilon(s) \prod_{i=1}^n \langle e_{s(i)}, \mu_i e_i \rangle \quad (18.272b)$$

$$= \sum_{s \in S_n} \epsilon(s) \prod_{i=1}^n \mu_i \langle e_{s(i)}, e_i \rangle. \quad (18.272c)$$

Vu que la base est orthogonale,  $\langle e_{s(i)}, e_i \rangle = \delta_{s(i), i}$  et il ne reste, dans la somme sur  $S_n$ , que le terme  $s = \text{Id}$  dont la signature est  $\epsilon(\text{Id}) = 1$ . Donc

$$\det_B(f(B)) = \prod_{i=1}^n \mu_i = -1. \quad (18.273)$$

Et si  $H$  ne passe pas par l'origine ? Soit  $v \in \mathbb{R}^n$  tel que  $\tau_v(H)$  passe par l'origine (prendre pour  $v$  l'opposé de n'importe quel élément de  $H$ ). Le lemme 18.91 nous indique que

$$\sigma_{\tau_v(H)} = \tau_v \circ \sigma_H \circ \tau_v^{-1}. \quad (18.274)$$

Comme  $\epsilon$  est un morphisme,

$$\epsilon(\sigma_{\tau_v(H)}) = \epsilon(\tau_v \sigma_H \tau_v^{-1}) = \epsilon(\tau_v) \epsilon(\sigma_H) \epsilon(\tau_v^{-1}) = \epsilon(\sigma_H). \quad (18.275)$$

Nous avons utilisé le fait que  $\epsilon(g^{-1}) = \epsilon(g)^{-1}$ . Tout ça pour dire que

$$\epsilon(\sigma_H) = \epsilon(\sigma_{\tau_v(H)}) = -1 \quad (18.276)$$

parce que  $\tau_v(H)$  passe par l'origine. □

THOooQEW RooYeOIfZ

**Théorème 18.96** ([605]).

Une isométrie de  $(\mathbb{R}^n, d)$  préserve l'orientation<sup>26</sup> si et seulement si elle est composition d'un nombre pair de réflexions hyperplanes.

*Démonstration.* Le lemme 18.88 nous indique qu'une isométrie  $f$  de  $\mathbb{R}^n$  peut être décomposée en réflexions hyperplanes. Le lemme 18.95 nous dit que chacune de ces réflexions est négative. Donc si  $f = \sigma_1 \circ \dots \circ \sigma_r$ , alors

$$\epsilon(f) = \epsilon(\sigma_1) \dots \epsilon(\sigma_r) = (-1)^r. \quad (18.277)$$

□

## 18.9 Groupes finis d'isométries

DEFooCUYLooAlbtzv

**Définition 18.97.**

Si  $X$  est une partie finie de  $\mathbb{R}^n$ , le **barycentre** de  $X$  est le point

$$B_X = \frac{1}{|X|} \sum_{x \in X} x \quad (18.278)$$

où  $|X|$  est le cardinal de  $X$ .

---

<sup>26</sup>. Définition 18.93.

Cela est à mettre en relation avec la définition dans le cadre affine 11.30.

LEMooSEZYooYceLIb

**Lemme 18.98** ([605]).

Les applications affines de  $\mathbb{R}^n$  préservent le barycentre<sup>27</sup> des parties finies.

*Démonstration.* Soit une partie finie  $X$  de  $\mathbb{R}^n$  et une application affine  $f \in \text{Aff}(\mathbb{R}^n)$ . Nous devons prouver que

$$f(B_X) = B_{f(X)}. \quad (18.279)$$

Nous savons que toute application affine est une composée de translation et d'une application linéaire :  $f = \tau_v \circ g$  avec  $v \in \mathbb{R}^n$  et  $g \in \text{End}(n, \mathbb{R})$ . Nous vérifions le résultat séparément pour  $\tau_v$  et pour  $g$ .

D'une part,

$$B_{\tau_v(X)} = \frac{1}{|\tau_v(X)|} \sum_{y \in \tau_v(X)} y = \frac{1}{|X|} \sum_{x \in X} (x + v) = B_X + \frac{1}{|X|} \sum_{x \in X} v = B_X + v = \tau_v(B_X). \quad (18.280)$$

Nous avons utilisé le fait que  $X$  et  $\tau_v(X)$  possèdent le même nombre d'éléments, ainsi que le fait d'avoir une somme de  $|X|$  termes tous égaux à  $v$ .

D'autre part,

$$B_{g(X)} = \frac{1}{|X|} \sum_{x \in X} g(x) = g\left(\frac{1}{|X|} \sum_{x \in X} x\right) = g(B_X) \quad (18.281)$$

où nous avons utilisé la linéarité de  $g$  dans tous ses retranchements.  $\square$

PROFoooLAEBooWdcBoe

**Proposition 18.99.**

Points fixes d'un sous-groupe.

- (1) Soit  $H$  un sous-groupe fini des isométries de  $(\mathbb{R}^n, d)$ . Alors il existe  $v \in \mathbb{R}^n$  tel que  $f(v) = v$  pour tout  $f \in H$ .
- (2) Si  $H$  est un sous-groupe de  $\text{Isom}(\mathbb{R}^n, d)$  n'acceptant pas de point fixe, alors il est infini.

*Démonstration.* Le groupe  $H$  agit sur  $\mathbb{R}^n$ , et si  $x \in \mathbb{R}^n$  nous pouvons considérer son orbite

$$Hx = \{f(x) \text{ tel que } f \in H\}, \quad (18.282)$$

qui est une partie finie de  $\mathbb{R}^n$ . Considérons son barycentre  $v = B_{Hx}$ . Soit  $f \in H$ . Alors

$$f(v) = f(B_{Hx}) \quad (18.283a)$$

$$= B_{f(Hx)} \quad (18.283b)$$

$$= B_{Hx} \quad (18.283c)$$

$$= v, \quad (18.283d)$$

SUBEQoooQPBZoooYllbgN

SUBEQoooXWEGoooSoozYg

Justifications :

- Pour (18.283b), c'est le lemme 18.98.
- Pour (18.283c), c'est le fait que,  $f \in H$  étant donné, l'application  $g \mapsto fg$  est une bijection de  $H$ , donc

$$f(Hx) = \{(fg)(x) \text{ tel que } h \in H\} = \{g(x) \text{ tel que } g \in H\} = Hx. \quad (18.284)$$

Bref,  $v$  est fixé par  $H$ .

La seconde affirmation n'est rien d'autre que la contraposée de la première.  $\square$

PROFoooEUFlooDUIYzi

**Proposition 18.100.**

À propos de groupes finis d'isométries.

- (1) Tout sous-groupe fini de  $\text{Isom}(\mathbb{R}^n)$  est isomorphe à un sous-groupe fini de  $O(n)$ .

27. Définition 18.97.

(2) *Tout sous-groupe fini de  $\text{Isom}^+(\mathbb{R}^n)$  est isomorphe à un sous-groupe fini de  $\text{SO}(n)$ .*

*Démonstration.* Soit  $H$  un sous-groupe fini de  $\text{Isom}(\mathbb{R}^n)$  et  $v \in \mathbb{R}^n$  un élément fixé par  $H$  (comme garanti par la proposition 18.99). Nous posons

$$\begin{aligned}\phi: H &\rightarrow \text{Isom}(\mathbb{R}^n) \\ f &\mapsto \tau_v^{-1} \circ f \circ \tau_v.\end{aligned}\tag{18.285}$$

(1)  $\phi$  est un morphisme Les opération du type  $\phi = \text{Ad}(\tau_v)$  sont toujours des morphismes.

(2)  $\phi$  consiste à extraire la partie linéaire Si  $f = \tau_w \circ g$  alors

$$\phi(f)(x) = (\tau_{-v} \circ \tau_w \circ g \circ \tau_v)(x) \tag{18.286a}$$

$$= \tau_{w-v}(g(x) + g(v)) \tag{18.286b}$$

$$= g(x) + g(v) - v + w \tag{18.286c}$$

Mais  $g(v) + w = f(v)$  et nous savons que  $f(v) = v$ . Donc il ne reste que  $\phi(f)(x) = g(x)$ .

(3)  $\phi$  est injective Si  $f = \tau_w \circ g$  vérifie  $\phi(f) = \text{Id}$ , il faut en particulier que  $g = \text{Id}$ . Mais  $H$  est fini et ne peut donc pas contenir de translations non triviales. Donc  $w = 0$  et  $f = \text{Id}$ .

Donc  $\phi$  est une injection à valeur dans les transformations linéaires de  $\text{Isom}(\mathbb{R}^n)$ . Autrement dit,  $\phi$  est un isomorphisme entre  $H$  et son image, laquelle image est dans  $\text{O}(n)$ .

En ce qui concerne la seconde partie, si  $f \in \text{Isom}^+(\mathbb{R}^n)$ , alors  $\phi(f)$ , qui est la partie linéaire de  $f$  (théorème 18.80(5)), est dans  $\text{O}(n)$  avec un déterminant égal à 1. Autrement dit,  $\phi(f) \in \text{SO}(n)$ .  $\square$

L'extraction de la partie linéaire est injective ? Certe c'est prouvé, mais on peut se demander ce qu'il se passe si  $H$  contient deux éléments qui ont la même partie linéaire. Cela n'est pas possible parce si  $f_1 = \tau_{w_1} \circ g$  et  $f_2 = \tau_{w_2} \circ g$  sont dans  $H$  alors  $f_1 f_2^{-1} = \tau_{w_1 - w_2}$  est également dans  $H$ . Vu que  $H$  est fini, cela n'est possible que si  $w_1 = w_2$ , c'est-à-dire si  $f_1 = f_2$ .

### 18.9.1 Points fixés par une affinité

LEMooGUEGooTUXRsQ

**Lemme 18.101** ([607]).

Si  $n \geq 3$ , alors toute droite est intersection de deux plans non isotropes<sup>28</sup>.

*Démonstration.* Soit une droite  $d$  dans  $E$ , ainsi que  $x \in d$ . Soit  $x$  est isotrope, soit il ne l'est pas.

(1) Si  $x$  est isotrope

Étant donné que  $f$  est non dégénérée, il existe  $y \in E$  tel que  $f(x, y) \neq 0$ . Le plan  $P = \text{Span}\{x, y\}$  est alors non isotrope. En effet, considérons  $z \in P \cap P^\perp$ . Nous avons alors  $f(z, x) = f(z, y) = 0$ . Vu que  $z \in P$  nous pouvons l'écrire sous la forme  $z = \alpha x + \beta y$ . La première condition sur  $z$  donne :

$$0 = f(\alpha x + \beta y, x) = \beta f(y, x) \tag{18.287}$$

parce que  $f(x, x) = 0$ . Étant donné que  $f(x, y) \neq 0$  nous en déduisons que  $\beta = 0$ . La seconde condition donne alors  $f(\alpha x, y) = 0$ , ce qui donne  $\alpha = 0$ .

Le plan  $P$  étant non isotrope, le plan  $P^\perp$  est également non isotrope par le lemme 8.142.

Nous fixons maintenant  $z \in P^\perp$ .

L'élément  $z$  n'est pas dans  $P$  parce que  $P \cap P^\perp = \{0\}$ . L'élément  $y + z$  n'est pas dans  $P$  non plus parce que  $P$  est vectoriel,  $y \in P$  et  $z \notin P$ .

Nous considérons donc le plan  $Q = \text{Span}\{x, y + z\}$ .

Ce plan est non isotrope. En effet, soit  $u \in Q \cap Q^\perp$ . Nous avons  $u = \alpha x + \beta(y + z)$ , et cela doit vérifier  $f(u, x) = f(y, y + z) = 0$ . Nous avons d'abord

$$0 = f(u, x) = \underbrace{\alpha f(x, x)}_{=0} + \beta \underbrace{f(x, y)}_{\neq 0} + \beta \underbrace{f(x, z)}_{=0}. \tag{18.288}$$

28. Sous-espace isotrope, définition 8.140.

Nous en déduisons que  $\beta = 0$ . Il reste  $u = \alpha x$ , et nous imposons la seconde condition :

$$0 = f(u, y + z) = \alpha f(x, y) + \alpha f(x, z). \quad (18.289)$$

Vu que  $f(x, y) \neq 0$  et  $f(x, z) = 0$  nous déduisons  $\alpha = 0$  et donc  $u = 0$ .

Au final, les plans  $P$  et  $Q$  sont non isotropes et  $P \cap Q = d$ .

- (2) **Si  $x$  n'est pas isotrope** Nous supposons maintenant que  $x$  n'est pas isotrope. Le lemme 8.133 sur les dimensions, appliqué à l'espace de dimension 1 engendré par  $x$  donne  $\dim(x^\perp) + 1 = \dim(E)$ . Vu que  $\dim(E) \geq 3$ , nous déduisons que  $\dim(x^\perp) \geq 2$ .

Nous considérons une base  $f$ -orthogonale de  $x^\perp$  (proposition 8.242). Les vecteurs de cette base ne sont pas isotropes. Soient  $y$  et  $z$  les vecteurs de cette base. Il nous reste à prouver que les plans  $P = \text{Span}\{x, y\}$  et  $Q = \text{Span}\{x, z\}$  font l'affaire.

Vu que  $x$  n'est pas isotrope, les vecteurs  $y$  et  $z$  ne sont pas dans  $d$  ; ce sont donc de vrais plans de dimension 2, et leur intersection est  $d$ . Montrons que  $P$  n'est pas isotrope, en considérant  $v = \alpha x + \beta y \in P \cap P^\perp$ . Vu que  $v \in P^\perp$ , il vérifie les deux conditions  $f(v, x) = f(v, y) = 0$ .

La première condition donne

$$0 = \alpha \underbrace{f(x, x)}_{\neq 0} + \beta \underbrace{f(y, x)}_{=0} = \alpha f(x, x) \quad (18.290)$$

parce que  $y \in x^\perp$ . Il nous reste  $v = \beta y$ . Appliquons la seconde condition :

$$0 = f(v, y) = \alpha f(y, y). \quad (18.291)$$

Le scalaire  $f(y, y) \in \mathbb{K}$  n'est pas nul parce que nous avons choisi une base orthogonale. Au final  $v = 0$ , et  $P$  est non isotrope. Exactement le même raisonnement montre que  $Q$  n'est pas isotrope non plus.

□

PROFoooVEEUooJQmmkN

### Proposition 18.102 ([605]).

Si une isométrie de  $\mathbb{R}^n$  fixe un ensemble  $F$  de points, alors elle fixe l'espace affine engendré par  $F$ .

*Démonstration.* Soit  $f \in \text{Isom}(\mathbb{R}^n)$  fixant  $F$ . Par le théorème 8.154, c'est une application affine et l'ensemble  $\text{Fix}(f)$  des points fixés par  $f$  est un sous-espace affine de  $\mathbb{R}^n$ , grâce à la proposition 11.61.

Donc  $\text{Fix}(f)$  est un espace affine contenant  $F$ . Puisque l'espace affine engendré par  $F$  est l'intersection de tous les espaces affines contenant  $F$ , il est en particulier contenu dans  $\text{Fix}(f)$ . □

CORooZHZZooDgTzsW

### Corolaire 18.103.

Si  $f$  et  $g$  sont des isométries de  $\mathbb{R}^n$  qui coïncident sur  $F$ , alors elles coïncident sur l'espace affine engendré par  $F$ .

*Démonstration.* Nous considérons  $h = g^{-1} \circ f$  qui est une isométrie de  $\mathbb{R}^n$  fixant  $F$ . Elle fixe donc, par la proposition 18.102, l'espace affine engendré par  $F$ . Or tout point fixé par  $h$  est un point sur lequel  $g$  et  $f$  coïncident. □

## 18.10 Classification des isométries dans $\mathbb{R}^2$

### 18.10.1 Projection orthogonale

Une version pour les espaces de Hilbert sera le théorème 25.7

DEFooHXJTooNPYDFQ

### Proposition-Définition 18.104.

Soient un espace euclidien de dimension finie  $(E, \cdot)$ , un sous-espace vectoriel  $H_0$  ainsi que  $a, d \in E$ . Nous considérons le sous-espace affine  $H = H_0 + d$ .

Il existe un unique  $y \in H$  tel que

$$\|a - y\| = \inf\{\|a - x\| \text{ tel que } x \in H\}. \quad (18.292)$$

Ce  $y$  est la **projection orthogonale** de  $a$  sur  $H$  et est notée  $\text{proj}_H(a)$ .

LEMooXZM FooOPTjNx

### Lemme 18.105.

Soient un espace vectoriel euclidien  $(E, \cdot)$  ainsi que  $a \in E$ . Nous considérons un sous-espace affine  $H$  de  $E$ . Si  $h \in H$ , alors

$$a \cdot h = \text{proj}_H(a) \cdot h. \quad (18.293)$$

## 18.10.2 Réflexions

Soit un espace vectoriel  $E$  de dimension 2 muni d'un produit scalaire<sup>29</sup>. Cela pourrait très bien être  $\mathbb{R}^2$ , mais nous allons nous efforcer de l'appeler  $E$  pour rester un peu général.

DEFooLJKDooUaamen

### Lemme-Définition 18.106 (Caractérisation des réflexions).

Soit une droite  $\ell$  de  $\mathbb{R}^2$ . Il existe une unique application  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  telle que

- (1)  $f(x) = x$  pour tout  $x \in \ell$ .
- (2)  $f$  échange les côtés de  $\ell$ .
- (3)  $f$  laisse invariants les droites perpendiculaires à  $\ell$  et les cercles dont le centre est sur  $\ell$ .

Cette application est la **réflexion** d'axe  $\ell$ .

*Démonstration.* Soit  $x$  hors de  $\ell$  et  $p$  la droite perpendiculaire à  $\ell$  et passant par  $x$ . Nous avons  $f(x) \in p$ . En nommant  $P$  l'intersection entre  $\ell$  et  $p$ , nous considérons le cercle  $S(P, \|Px\|)$  qui est un cercle dont le centre est sur  $\ell$ . Il contient  $x$  et donc  $f(x) \in S(P, \|Px\|)$ .

Donc  $f(x) \in p \cap S(P, \|Px\|)$ . L'intersection entre un cercle et une droite contient de façon générique deux points. L'un est  $x$ , mais  $f(x) = x$  n'est pas possible parce que  $x$  est hors de  $\ell$  et  $f$  doit inverser les côtés de  $\ell$ . Donc  $f(x)$  est l'autre.

Cela prouve l'unicité. En ce qui concerne l'existence, il suffit de noter que la réflexion  $\sigma_\ell$  satisfait les contraintes.  $\square$

LEMooZSDRooUkNYer

### Lemme 18.107.

Soit une droite  $\ell$  et  $A \in E$ . Alors

$$\sigma_\ell(A) = 2\text{proj}_\ell(A) - A \quad (18.294)$$

EQooVUQDooKuwszl

où  $\text{proj}_\ell$  est l'opération de projection orthogonale sur la droite  $\ell$ .

*Démonstration.* Nous posons  $f(X) = 2\text{proj}_\ell(X) - X$  et nous allons montrer que  $f = \sigma_\ell$  en vérifiant les conditions de la définition 18.106. Nous nous gardons bien de faire un raisonnement du type « nous allons montrer que  $f$  et  $\sigma_\ell$  coïncident sur deux points, et sont donc égales par le corolaire 18.103 » parce que nous ne savons pas encore que  $\sigma_\ell$  est une application affine, ni même que c'est une isométrie.

Si  $X \in \ell$  alors  $\text{proj}_\ell(X) = X$  et nous avons  $f(X) = 2X - X = X$ . Donc  $\ell$  est conservée.

En ce qui concerne les deux côtés de  $\ell$ , il existe une application linéaire  $s: E \rightarrow \mathbb{R}$  et une constante  $c \in \mathbb{R}$  telles qu'en posant  $\ell(X) = s(X) + c$ , la droite  $\ell$  soit le lieu des points  $X$  tels que  $\ell(X) = 0$ . Un côté de la droite est  $\ell < 0$  et l'autre côté est  $\ell > 0$ . Nous avons :

$$\ell(f(A)) = \ell(2\text{proj}_\ell(A) - A) \quad (18.295a)$$

$$= s(2\text{proj}_\ell(A) - A) + c \quad (18.295b)$$

$$= 2s(\text{proj}_\ell(A)) - s(A) + c \quad (18.295c)$$

$$= s(\text{proj}_\ell(A)) - s(A) \quad (18.295d)$$

$$= -c - s(A) \quad (18.295e)$$

$$= -\ell(A) \quad (18.295f)$$

29. Définition 8.164.

où nous avons utilisé le fait que,  $\text{proj}_\ell(A)$  étant sur  $\ell$ ,  $s(\text{proj}_\ell(A)) + c = 0$ . Nous avons donc  $\ell(f(A)) = -\ell(A)$ , ce qui indique que  $A$  et  $f(A)$  sont de part et d'autre de  $\ell$ .

Si  $d$  est une droite perpendiculaire à  $\ell$  et si  $A \in d$  alors  $f(A) = 2\text{proj}_\ell(A) - A = (\text{proj}_\ell(A) - A) + A \in d$  parce que  $\text{proj}_\ell(A) \in d$  du fait que  $d$  soit précisément perpendiculaire à  $\ell$ . Nous avons aussi utilisé le fait que si  $A, B, C \in d$  alors  $(B - A) + C \in d$ ; pensez que  $B - A$  est un vecteur directeur et que  $C$  est un point de  $d$ .

Enfin soit  $K \in \ell$  et un cercle  $S(K, r)$  centré en  $K$ . Soit  $A \in S(K, r)$ ; nous devons vérifier que  $f(A) \in S(K, r)$ . Le segment  $[A, f(A)]$  est par définition perpendiculaire à  $\ell$ . Soit  $M$ , le milieu, qui est sur la droite  $\ell$ . Les triangles  $AMK$  et  $f(A)MK$  sont rectangles en  $M$ , et  $\|AM\| = \|Mf(A)\|$ . Le théorème de Pythagore donne  $\|AK\| = \|f(A)K\|$ . Donc le cercle centré en  $K$  est alors préservé par  $f$ .

Nous en déduisons que  $f = \sigma_\ell$ . □

LEMooNYIWooXanBXh

### Lemme 18.108.

Soit une droite  $\ell$  dans  $\mathbb{R}^2$  ainsi que  $a \in \mathbb{R}^2$ . Nous avons

$$\text{proj}_\ell(\sigma_\ell(a)) = \text{proj}_\ell(a). \quad (18.296)$$

PROPOoFSVEooWmJsnv

### Proposition 18.109 ([608]).

Une réflexion est une isométrie de  $(E, d)$  où  $d(A, B) = \|A - B\|$ .

*Démonstration.* Soient  $A, B \in E$ ; il faut vérifier que  $\|A - B\| = \|\sigma_\ell(A) - \sigma_\ell(B)\|$ . Pour cela nous écrivons

$$B - A = \underbrace{B - \text{proj}_\ell(B)}_{=a} + \underbrace{\text{proj}_\ell(B) - \text{proj}_\ell(A)}_{=b} + \underbrace{\text{proj}_\ell(A) - A}_{=c}. \quad (18.297)$$

Vu que  $b \perp a$  et  $b \perp c$  (lemme 18.105) nous avons

$$\|B - A\|^2 = \langle B - A, B - A \rangle = \|a\|^2 + 2\langle a, c \rangle + \|b\|^2 + \|c\|^2. \quad (18.298)$$

Nous pouvons jouer le même jeu avec  $\sigma_\ell(B) - \sigma_\ell(A)$  en tenant compte du lemme 18.108 et que

$$\sigma_\ell(A) - \text{proj}_\ell(A) = 2\text{proj}_\ell(A) - A - \text{proj}_\ell(A) = -(A - \text{proj}_\ell(A)). \quad (18.299)$$

Là nous avons utilisé le lemme 18.107. Ce que nous trouvons est que

$$\sigma_\ell(B) - \sigma_\ell(A) = -a + b - c, \quad (18.300)$$

et donc encore une fois

$$\|\sigma_\ell(B) - \sigma_\ell(A)\|^2 = \|a\|^2 + 2\langle a, c \rangle + \|b\|^2 + \|c\|^2. \quad (18.301)$$

□

### Remarque 18.110.

Il faut bien comprendre que si l'axe de la réflexion ne passe pas par 0 (le zéro de l'espace vectoriel normé  $(E, \|\cdot\|)$ ), la réflexion n'est pas une isométrie de  $(E, \|\cdot\|)$  au sens où nous n'avons pas  $\|\sigma_\ell(x)\| = \|x\|$ .

LEMooTCIEooXdyuHu

### Lemme 18.111.

Si  $A'$  est l'image de  $A$  par  $\sigma_\ell$  alors  $\ell$  est la médiatrice du segment  $[A, A']$ .

*Démonstration.* Soit  $M \in \ell$ . Nous avons

$$\|A - M\|^2 = \|\text{proj}_\ell(A) - A\|^2 + \|\text{proj}_\ell(A) - M\|^2 \quad (18.302)$$

parce que  $A - \text{proj}_\ell(A) \perp M - \text{proj}_\ell(A)$ . Par ailleurs, vu que  $\sigma_\ell(A) = 2\text{proj}_\ell(A) - A$  et que  $\text{proj}_\ell(A) = \text{proj}_\ell(A')$ ,

$$\|\text{proj}_\ell(A) - A\| = \|\text{proj}_\ell(A') - A'\|. \quad (18.303)$$

Nous avons donc

$$\|\sigma_\ell(A) - M\|^2 = \|A - M\|^2, \quad (18.304)$$

ce qui prouve que  $M$  est sur la médiatrice de  $[A', A]$  par le lemme 12.158. □

**18.112.**

Si  $l$  est une droite dans  $\mathbb{R}^2$ , nous avons la réflexion  $\sigma_l \in \text{Isom}(\mathbb{R}^2)$  d'axe  $l$ . Cela est une isométrie et donc une application affine par le théorème 8.154. Le lemme suivant détermine comment la réflexion  $\sigma_\ell$  se décompose en une translation et une application linéaire.

LEMooVOJLooCFgdNG

**Lemme 18.113.**

Soit une droite  $\ell$ . Alors

$$\sigma_\ell = \tau_{2w} \circ \sigma_{\ell_0} \quad (18.305)$$

où  $\ell_0$  est la droite parallèle à  $\ell$  passant par l'origine, et  $w$  est le vecteur perpendiculaire à  $\ell$  tel que  $\ell_0 = \ell + w$ .

*Démonstration.* Il faut trouver trois points non alignés sur lesquels les deux applications coïncident ; cela suffira par le corolaire 18.103.

Pour tous les points de  $\ell_0$ , l'égalité fonctionne parce que si  $x \in \ell_0$ ,

$$\sigma_\ell(x) = x + 2w, \quad (18.306)$$

tandis que

$$\sigma_{\ell_0}(x) + 2w = x + 2w \quad (18.307)$$

du fait que  $\sigma_{\ell_0}(x) = x$ .

Si  $x \in \ell$ , alors

$$\sigma_\ell(x) = x \quad (18.308)$$

tandis que

$$\sigma_{\ell_0}(x) + 2w = x - 2w + 2w = x. \quad (18.309)$$

Donc les applications affines  $\sigma_\ell$  et  $x \mapsto \sigma_{\ell_0}(x) + 2w$  coïncident sur  $\ell$  et  $\ell_0$ . Elles coïncident donc partout.  $\square$

LEMooRVKEooSHrIR

**Lemme 18.114 ([605]).**

Une isométrie de  $(\mathbb{R}^2, d)$  fixant tous les points d'une droite est soit l'identité soit la réflexion de cette droite.

*Démonstration.* Soit  $l$  une droite fixée par  $f$ , et soient  $x, y \in l$  et  $z \notin l$  (avec  $x \neq y$ ). Le fait que  $x$  et  $y$  soient des points fixes de  $f$  implique

$$\begin{cases} d(x, f(z)) = d(x, z) \\ d(y, f(z)) = d(y, z) \end{cases} \quad (18.310a)$$

$$(18.310b)$$

ce qui signifie que  $f(z)$  est sur l'intersection des deux cercles<sup>30</sup>  $S(x, d(x, z))$  et  $S(y, d(y, z))$ , et comme ce sont deux cercles centrés sur la droite  $l$ , les intersections sont liées par  $\sigma_l$ . Autrement dit, les intersections sont  $z$  et  $\sigma_l(z)$ .

Si  $f(z) = z$  alors  $f$  fixe trois points non alignés et fixe donc  $\mathbb{R}^2$ , c'est-à-dire  $f = \text{Id}$ .

Si par contre  $f(z) = \sigma_l(z)$  alors les isométries  $f$  et  $\sigma_l$  coïncident sur trois points et coïncident donc partout par le corolaire 18.103 :  $f = \sigma_l$ .  $\square$

**18.10.3 Translations et réflexions**

LEMooMKTXooYKZcdQ

**Lemme 18.115.**

Si  $A \in \mathbb{R}^2$ , si  $\ell$  est une droite de  $\mathbb{R}^2$ , alors nous avons

$$\text{proj}_{\tau_A(\ell)} = \tau_A \circ \text{proj}_\ell \circ \tau_A^{-1}. \quad (18.311)$$

30. L'intersection existe parce que  $d(x, z) + d(y, z) > d(x, y)$ .

*Démonstration.* Soit  $x \in \mathbb{R}^2$ . Soit  $v$  unitaire dans la direction<sup>31</sup> de  $\ell$ . La condition  $q = \text{proj}_\ell(x)$  est le système

$$\begin{cases} q \in \ell \\ (q - x) \cdot v = 0. \end{cases} \quad \begin{array}{l} (18.312a) \\ (18.312b) \end{array}$$

Nous voulons prouver que  $\tau_A(q) = \text{proj}_{\tau_A(\ell)}(\tau_A(x))$ , c'est-à-dire que

$$\begin{cases} \tau_A(q) \in \tau_A(\ell) \\ (\tau_A(q) - \tau_A(x)) \cdot v = 0. \end{cases} \quad \begin{array}{l} (18.313a) \\ (18.313b) \end{array}$$

Nous avons utilisé le fait que  $v$  est un vecteur unitaire dans la direction de  $\tau_A(\ell)$  aussi bien que de  $\ell$ .

Puisque  $q \in \ell$ , il est bien entendu que  $\tau_A(q) \in \tau_A(\ell)$ . D'autre part,  $\tau_A(q) - \tau_A(x) = q - x$ , donc

$$(\tau_A(q) - \tau_A(x)) \cdot v = (q - x) \cdot v = 0. \quad (18.314)$$

□

LEMooSMMMooAqsHWb

### Lemme 18.116.

Si  $\ell$  est une droite de  $\mathbb{R}^2$ , si  $A \in \mathbb{R}^2$ , alors

$$\sigma_{\tau_A(\ell)} = \tau_A \sigma_\ell \tau_A^{-1}. \quad (18.315)$$

*Démonstration.* Il s'agit d'un calcul mettant en scène les lemmes 18.107 et 18.115 :

$$(\sigma_{\tau_A(\ell)} \tau_A)(x) = 2 \text{proj}_{\tau_A(\ell)}(\tau_A(x)) \quad (18.316a)$$

$$= 2\tau_A(\text{proj}_\ell(x)) - \tau_A(x) \quad (18.316b)$$

$$= 2 \text{proj}_\ell(x) + 2A - x - A \quad (18.316c)$$

$$= 2 \text{proj}_\ell(x) - x + A \quad (18.316d)$$

$$= \tau_A(\sigma_\ell(x)). \quad (18.316e)$$

□

## 18.10.4 Rotations

DEFooFUBYooHGXPbm

**Définition 18.117** (Rotation en dimension 2).

Une **rotation** d'un espace euclidien de dimension 2 est une composée de deux réflexions d'axes non parallèles. L'identité est une rotation.

### 18.118.

Quelques remarques à propos de cette définition.

- (1) Attention : nous ne parlons pas encore de rotations « vectorielles » : ici le centre de la rotation (que nous n'avons pas encore défini) peut ne pas être 0.
- (2) Dans la même veine : plus tard, lorsque nous saurons que les rotations sont des isométries de  $(E, d)$  où  $d(X, Y) = \|X - Y\|$ , nous allons en réalité beaucoup plus souvent parler de rotations centrées en l'origine qu'en un point quelconque. C'est pourquoi à partir de 18.129 nous dirons le plus souvent « rotation » pour « rotation centrée en 0 ». D'où les énoncés comme « les rotations sont les matrice orthogonales » (corolaire 18.137), qui *stricto sensu* de la définition 18.117, sont faux.
- (3) Une rotation est composée de deux réflexions d'axes non parallèles. Il est cependant trop tôt pour décréter que l'intersection de ces axes est le centre de la rotation. Rien ne dit en effet pour l'instant que deux décompositions différentes de la même rotation, avec des axes différents donnent le même point d'intersection.

31. Cela signifie qu'il existe  $a \in \mathbb{R}^2$  tel que  $\ell = a + \mathbb{R}v$ .



- (4) Pourquoi ajouter l'identité? Pour avoir un groupe. Dans le cas vectoriel, il est suffisant de demander d'être une composée de deux réflexions, parce que toutes les réflexions vectorielles ont des axes qui s'intersectent en 0. Le cas des axes parallèles est seulement le cas des axes confondus et revient à l'identité.

Si nous voulons avoir un groupe même pour les rotations centrées ailleurs qu'en zéro, nous devons ajouter « à la main » l'identité.

Toutes ces remarques se résument par : « tout devient compliqué du fait que nous voulons considérer également les rotations centrées ailleurs qu'en zéro ». En se contentant du cas vectoriel, de nombreuses choses sont plus simples.

CORooNKKIooPGOUJI

### Corolaire 18.119.

Si  $A \neq B$  dans  $E$  alors il existe une unique réflexion envoyant  $A$  sur  $B$ .

*Démonstration.* En ce qui concerne l'existence, la réflexion dont l'axe est la médiatrice de  $[A, B]$  fait l'affaire. En ce qui concerne l'unicité, le lemme 18.111 nous dit que si  $A$  est envoyé sur  $B$ , l'axe est forcément la médiatrice de  $[A, B]$ .  $\square$

LEMooIJELOoLWqBfE

### Lemme-Définition 18.120 ([609]).

Soit une rotation  $r = \sigma_1 \circ \sigma_2$  différente de l'identité.

- (1) Elle admet un unique point fixe.
- (2) Ce point fixe est l'intersection des axes  $\ell_1 \cap \ell_2$ .

Le **centre** d'une rotation (autre que l'identité) est cet unique point fixe.

*Démonstration.* Nous nommons  $O = \ell_1 \cap \ell_2$ . Soit  $A \in E$ , et supposons que  $r(A) = A$ . Nous avons  $\sigma_1 \circ r = \sigma_2$  et donc

$$\sigma_1(A) = (\sigma_1 \circ r)(A) = \sigma_2(A). \quad (18.317)$$

On pose  $B = \sigma_1(A)$ . Alors  $\sigma_1$  et  $\sigma_2$  envoient tous deux  $A$  sur  $B$ .

Si  $A = B$  alors  $A$  est fixé par  $\sigma_1$  et donc appartient à  $\ell_1$ . Même chose pour  $A$  qui est fixé par  $\sigma_2$ , et donc  $A \in \ell_2$ . Cela donne  $A = B = O$ , et alors, le point fixé par  $r$  est  $O$ .

Si  $A \neq B$  alors il existe une unique réflexion envoyant  $A$  sur  $B$  (corolaire 18.119). L'unicité signifie que  $\sigma_1 = \sigma_2$ . Dans ce cas,  $r = \sigma_1 \circ \sigma_2 = \text{Id}$ .  $\square$

NORMooDPBOooKkRuTn

### 18.121.

La rotation  $\sigma_1 \circ \sigma_2$  laisse évidemment fixé le point  $\ell_1 \cap \ell_2$ . Si  $\sigma_1 \circ \sigma_2 = \sigma_a \circ \sigma_b$  alors rien n'oblige les axes de  $\sigma_1$  et  $\sigma_2$  d'être identiques à ceux de  $\sigma_a$  et  $\sigma_b$ . Mais l'intersection  $\ell_1 \cap \ell_2$  doit être la même que l'intersection  $\ell_a \cap \ell_b$  parce que c'est l'unique point fixé par la composée. Cela nous permet de poser la définition suivante.

LEMooTZNWooTVOklu

### Lemme 18.122.

Les rotations sont des isométries pour la distance :  $\|X - Y\| = \|r(X) - r(Y)\|$ .

*Démonstration.* Si  $r = \sigma_1 \circ \sigma_2$ , en utilisant le fait que  $\sigma_1$  et  $\sigma_2$  sont des isométries de  $(E, d)$  (18.109) nous avons :

$$d(X, Y) = d(\sigma_2(X), \sigma_2(Y)) = d(\sigma_1\sigma_2(X), \sigma_1\sigma_2(Y)) = d(r(X), r(Y)). \quad (18.318)$$

 $\square$ 

Ce lemme nous dit qu'une rotation de centre  $O$  vérifie  $\|OX\| = \|Or(X)\|$  pour tout  $X$ .

PROPooNXJKooEDOcZh

### Proposition 18.123 ([1, 609]).

Soient  $A, B, O \in E$  tels que  $\|AO\| = \|BO\| \neq 0$ . Alors il existe une unique rotation  $r$  centrée en  $O$  telle que  $r(A) = B$ .

*Démonstration.* Existence et unicité séparément.

- (1) **Existence** Si  $A = B$ , l'identité fait l'affaire. Sinon,  $\|A - O\| = \|B - O\|$  implique que la médiatrice de  $[A, B]$  contient  $O$ . Soit  $\sigma_m$  la réflexion selon cette médiatrice. La rotation  $\sigma_m \circ \sigma_{(AO)}$  convient.
- (2) **Unicité** Soit  $r$  une rotation de centre  $O$  et telle que  $r(A) = B$ . Si  $A = B$  alors  $r = \text{Id}$  parce qu'une rotation autre que l'identité ne fixe que son centre par le lemme 18.120. Nous supposons que  $A \neq B$ .

Nous posons  $g = \sigma_m \circ r$ . Alors  $g(A) = \sigma_m(B) = A$  parce que  $\sigma_m(B) = A$  et  $r(A) = B$ . Cela signifie que  $g$  est une isométrie qui fixe  $A$ .

- (2a) **Si  $A$  et  $O$  ne sont pas alignés** Attention : ici  $O$  est un point de  $E$ , pas le zéro de l'espace vectoriel  $E$ . Lorsqu'on dit que  $A$  et  $O$  ne sont pas alignés, nous parlons bien d'alignement avec le zéro de  $E$ .

Nous avons  $g(A) = A$  et  $g(O) = O$ . Donc  $g$  coïncide avec  $\sigma_{(AO)}$  en deux points non alignés, c'est-à-dire en deux points pour lesquels l'espace engendré est tout  $E$ . Nous en déduisons que  $g = \sigma_{(AO)}$ .

- (2b) **Si  $A$  et  $O$  sont alignés** Soit maintenant un point  $C$  tel que  $A - O \perp C - O$  et

$$\|OC\| = \|OA\| = \|OB\|. \quad (18.319)$$

Puisque  $g$  est une isométrie pour la distance sur  $E$ , pas pour la norme, nous ne pouvons pas écrire  $g(C - O) \perp g(A - O)$  à partir de  $C - O \perp A - O$ . Nous décomposons  $g(X) = s(X) + G$  où  $s$  est linéaire sur  $E$ . Il est vite vu que  $s$  est une isométrie de  $(E, \|\cdot\|)$  :

$$\|X - Y\| = \|g(X) - g(Y)\| = \|s(X) + G - s(Y) - G\| = \|s(X) - s(Y)\| = \|s(X - Y)\| \quad (18.320)$$

pour tout  $X, Y \in E$ . Nous avons de plus  $g(A) = A$  et  $g(O) = O$ , ce qui donne  $O = s(O) + G$  et  $A = s(A) + G$ . En égalisant les valeurs de  $G$  nous avons

$$O - s(O) = A - s(A). \quad (18.321)$$

Comme  $s$  est une isométrie (une vraie) nous avons

$$s(A - O) \perp s(C - O), \quad (18.322)$$

mais  $s(A - O) = s(A) - s(O) = A - O$  par (18.321). Donc

$$A - O \perp s(C - O). \quad (18.323)$$

Nous en concluons que  $s(C - O) = \pm(C - O)$ . Parce que les vecteurs  $\pm(C - O)$  sont les deux seuls de norme  $\|AO\| = \|CO\|$  à être perpendiculaire à  $A - O$ . Rappel : la définition de  $C$ , et le fait que nous soyons en dimension 2.

Est-il possible d'avoir  $s(C - O) = C - O$  ? Cela donnerait

$$g(A - O) = s(A) - s(O) + G = s(A) - O + O - s(O) + G = A - O + G \quad (18.324a)$$

$$g(C - O) = C - O + G, \quad (18.324b)$$

ce qui signifierait que  $g$  et  $\tau_G$  coïncideraient sur les points  $A - O$  et  $C - O$ , et donc seraient égaux par le corolaire 18.103. Cela est cependant impossible parce que  $g$  fixe au moins les points  $A$  et  $O$  alors que la translation ne fixe aucun point. Nous en déduisons  $s(C - O) = -(C - O)$ .

Nous avons aussi, parce que  $(AO)$  est une droite passant par l'origine que

$$\sigma_{(AO)}(A - O) = A - O \quad (18.325)$$

et parce que  $C - O$  est perpendiculaire à cette droite :

$$\sigma_{(AO)}(C - O) = -(C - O). \quad (18.326)$$

Nous avons donc quand même que  $g$  et  $\sigma_{(AO)}$  coïncident sur deux points non alignés :  $A - O$  et  $C - O$ .

Dans tous le cas,  $g = \sigma_{(AO)}$ . Nous avons donc

$$\sigma_{(OA)} = \sigma_m \circ r, \quad (18.327)$$

et donc  $r$  est fixé à

$$r = \sigma_m \circ \sigma_{(OA)}. \quad (18.328)$$

□

### 18.124.

Anticipons un peu et faisons semblant de déjà connaître les matrices et les fonctions trigonométriques. La proposition 18.123 nous dit qu'il existe une seule rotation amenant  $A$  sur  $B$ . Vous pourriez objecter que le point  $(1, 0)$  peut être amené sur  $(0, -1)$  soit par la rotation d'angle  $3\pi/2$ , soit par celle d'angle  $-\pi/2$ . Il n'en est rien parce que ces deux rotations sont les mêmes ! Pensez-y. En tant qu'application  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ , la rotation  $R_{3\pi/2}$  est égale à  $R_{-\pi/2}$ .

Une rotation donnée peut être écrite de beaucoup de façons comme composée de deux réflexions. En fait d'autant de façons qu'il y a de réflexions.

PROPooKAZEooLTHWKe

### Proposition 18.125 ([609]).

Soit une rotation  $r$  de  $E$  centrée en  $O$ . Pour toute réflexion  $\sigma_\ell$  telle que le centre de  $r$  soit sur  $\ell$ , il existe une réflexion  $\sigma_1$  telle que  $r = \sigma_1 \circ \sigma_\ell$ . Il existe aussi une réflexion  $\sigma_2$  telle que  $r = \sigma_\ell \circ \sigma_2$ .

*Démonstration.* Si  $r = \text{Id}$  c'est bon avec  $\sigma_1 = \sigma_2 = \sigma_\ell$ . Sinon nous considérons  $A \neq O$  sur  $\ell$ , et  $B = r(A)$ . Nous savons que  $B \neq A$  parce que  $O$  est le seul point de  $E$  fixé par  $r$  (proposition 18.120). Il existe une réflexion (unique)  $\sigma_1$  faisant  $\sigma_1(A) = B$ , et c'est la réflexion dont l'axe est la médiatrice de  $[A, B]$ . Le point  $O$  est sur cette médiatrice parce que les rotations sont des isométries de  $(E, d)$  (lemme 18.122).

La rotation  $\sigma_1 \circ \sigma_\ell$  vérifie

$$(\sigma_1 \circ \sigma_\ell)(A) = \sigma_1(A) = B. \quad (18.329)$$

Or  $\|OA\| = \|OB\|$ , donc il y a unicité de la rotation centrée en  $O$  portant  $A$  sur  $B$  (proposition 18.123) ; nous avons donc  $r = \sigma_1 \circ \sigma_\ell$ .

En ce qui concerne  $r = \sigma_\ell \circ \sigma_2$ , il faut appliquer ce que nous venons de faire à la rotation  $r^{-1}$  : il existe  $\sigma_2$  tel que  $r^{-1} = \sigma_2 \circ \sigma_\ell$ , ce qui donne

$$r = \sigma_\ell \circ \sigma_2. \quad (18.330)$$

□

### 18.10.5 Rotation d'un angle donné

DEFooADTDooKIZbrw

#### Lemme-Définition 18.126.

Soit  $\theta \in \mathbb{R}$ . Nous considérons l'application linéaire  $R_0(\theta) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dont la matrice dans la base canonique est<sup>32</sup>

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (18.331)$$

ITEMooIEKJooZfsAui

(1) Nous avons

$$R_0(\theta) = \sigma_\ell \circ s \quad (18.332)$$

EQooEVCTooRpTDDq

où  $s$  est la réflexion d'axe horizontal et  $\ell$  est la droite

$$\ell = \mathbb{R} \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}. \quad (18.333)$$

ITEMooBEYOooMHRRYk

(2) L'application  $R_0(\theta)$  est une rotation autour de  $(0, 0)$ .

32. Pour la définition des fonctions trigonométriques, définition 18.1.

(3) Si  $A \in \mathbb{R}^2$ , alors l'application

$$R_A(\theta) = \tau_A \circ R_0(\theta) \circ \tau_A^{-1} \quad (18.334)$$

est une rotation autour de  $A$  nommée **rotation d'angle**  $\theta$ .

*Démonstration.* Nous allons prouver l'égalité (18.332) en calculant les deux membres sur les vecteurs  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  et  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  qui forment une base.

(1) **Pour**  $p = (1, 0)$  D'abord la chose facile<sup>33</sup> :  $s(1, 0) = (1, 0)$ .

Pour calculer  $\sigma_\ell(1, 0)$ , nous utilisons le lemme 18.107 ; nous commençons par chercher la projection orthogonale  $q$  de  $p = (1, 0)$  sur  $\ell$ . Nous posons

$$q = \lambda \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \quad (18.335)$$

et nous cherchons  $\lambda$  satisfaisant  $q \cdot (q - p) = 0$ . Un peu de calculs passant par (18.5) nous donne

$$q \cdot (q - p) = \lambda(\lambda - \cos(\theta/2)). \quad (18.336)$$

Les deux solutions sont  $\lambda = 0$  et  $\lambda = \cos(\theta/2)$ . Mais la solution  $\lambda = 0$  revient à dire que la droite  $\ell$  est verticale, c'est-à-dire  $\cos(\theta/2) = 0$ . Donc la solution est toujours donnée par

$$\lambda = \cos(\theta/2). \quad (18.337)$$

Nous introduisons cette valeur dans (18.335) pour fixer  $q$ , et nous utilisons la formule du lemme 18.107 :

$$\sigma_\ell \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2q - p = \begin{pmatrix} 2 \cos^2(\theta/2) - 1 \\ 2 \cos(\theta/2) \sin(\theta/2) \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}. \quad (18.338)$$

Nous avons utilisé les formules trigonométriques de duplication d'angle (corolaire 18.15).

(2) **Pour**  $p = (0, 1)$  Cette fois  $s(p) = (0, -1)$  et l'équation pour déterminer  $\lambda$  est

$$0 = q \cdot (q - p) = \begin{pmatrix} \lambda \cos(\theta/2) \\ \lambda \sin(\theta/2) \end{pmatrix} \cdot \begin{pmatrix} \lambda \cos(\theta/2) \\ \lambda \sin(\theta/2) - 1 \end{pmatrix}. \quad (18.339)$$

Nous trouvons  $\lambda(\lambda + \sin(\theta/2)) = 0$ . Le cas  $\lambda = 0$  signifie que la droite  $\ell$  est horizontale et donc que  $\sin(\theta/2) = 0$ . Donc la solution est dans tous les cas

$$\lambda = -\sin(\theta/2). \quad (18.340)$$

Nous trouvons

$$\sigma_\ell \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2q - p = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}. \quad (18.341)$$

Le calcul de  $\sigma_\ell \circ s$  étant effectué sur une base, il est facile de reconstituer la matrice

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (18.342)$$

Cette matrice étant celle, par définition, de  $R_0(\theta)$ , nous avons montré que  $R_0(\theta)$  était bien une rotation. Nous avons prouvé les points (1) et (2).

---

33. Ici comme partout dans le Frido nous ne faisons aucune différence entre  $(a, b)$  et  $\begin{pmatrix} a \\ b \end{pmatrix}$  ; ce sont seulement deux façons différentes d'écrire le même élément de  $\mathbb{R}^2$ . Nous ne faisons pas semblant de croire que l'un ou l'autre serait un « covecteur » suivant que l'on tourne notre page dans un sens ou un autre.

Nous passons maintenant au point (3). Il est facile de voir que  $A$  est un point fixe de  $R_A(\theta)$  parce que  $\tau_A^{-1}(A) = (0, 0)$ .

Nommons  $\ell'$  la droite horizontale  $\mathbb{R}(1, 0)$ . Nous avons, par le point précédent  $R_0(\theta) = \sigma_\ell \circ \sigma_{\ell'}$ . En introduisant astucieusement  $\tau_A^{-1}\tau_A$  dans l'expression définissant  $R_A(\theta)$ , nous avons

$$R_A(\theta) = \tau_A \sigma_\ell \sigma_{\ell'} \tau_A^{-1} = \underbrace{\tau_A \sigma_\ell \tau_A^{-1}}_{=\sigma_{\tau_A(\ell)}} \underbrace{\tau_A \sigma_{\ell'} \tau_A^{-1}}_{=\sigma_{\tau_A(\ell')}} = \sigma_{\tau_A(\ell)} \sigma_{\tau_A(\ell')}. \quad (18.343)$$

Nous avons utilisé le lemme 18.116.

Nous voyons que  $R_A(\theta)$  est une composée de deux réflexions se coupant en  $A$ . C'est donc une rotation centrée en  $A$ . □

### 18.10.6 Rotations vectorielles

L'expression « rotation vectorielle » signifie rotation centrée en zéro. Elles sont « vectorielles » parce qu'elles sont linéaires comme nous le voyons à présent.

#### Proposition 18.127.

Quelques résultats à propos de rotations.

- (1) Toutes les rotations de  $\mathbb{R}^2$  centrées en  $(0, 0)$  sont de la forme  $R_0(\theta)$  <sup>34</sup> pour un  $\theta \in \mathbb{R}$ .
- (2) Les rotations de  $\mathbb{R}^2$  centrées en  $(0, 0)$  sont des applications linéaires.
- (3) Si  $r$  est une rotation dans  $\mathbb{R}^2$ , il existe  $A \in \mathbb{R}^2$  et  $\theta \in \mathbb{R}$

$$r = \tau_A \circ R_0(\theta) \circ \tau_A^{-1}. \quad (18.344)$$

*Démonstration.* Soit  $r$  une rotation centrée en  $(0, 0)$ . Par la proposition 18.125, il existe une droite  $\ell$  passant par  $(0, 0)$  telle que  $r = \sigma_\ell \circ s$  où  $s$  est la réflexion d'axe horizontal :  $s(x, y) = (x, -y)$ .

Soit un vecteur unitaire  $v \in \ell$ . Puisque  $v \in S^1$ , la proposition 18.59 nous donne  $t \in [0, 2\pi[$  tel que

$$v = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}. \quad (18.345)$$

En posant  $\theta = 2t$  nous avons

$$\ell = \mathbb{R} \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}, \quad (18.346)$$

et donc  $r = R_0(\theta)$  qui est l'application linéaire définie en 18.126. Ceci prouve les points (1) et (2).

Et enfin nous voyons le point (3). Soit  $A$  l'unique point fixe de la rotation  $r$ . Cette dernière s'écrit alors  $r = \sigma_{\ell_1} \circ \sigma_{\ell_2}$  où  $\ell_1$  et  $\ell_2$  sont des droites telles que  $\ell_1 \cap \ell_2 = \{A\}$ .

En utilisant le lemme 18.116, nous avons  $\sigma_{\ell_i} = \tau_A \circ \sigma_{\tau_A^{-1}(\ell_i)} \circ \tau_A^{-1}$ . En substituant, et en nous rendant compte que  $\tau_A^{-1}\tau_A = \text{Id}$ , nous avons

$$r = \sigma_{\ell_1} \circ \sigma_{\ell_2} = \tau_A r_0 \tau_A^{-1} \quad (18.347)$$

où  $r_0$  est la rotation  $\sigma_{\tau_A^{-1}(\ell_1)} \circ \sigma_{\tau_A^{-1}(\ell_2)}$ . Cette dernière est une rotation autour de  $(0, 0)$  parce que  $\tau_A^{-1}(\ell_1) \cap \tau_A^{-1}(\ell_2) = \{0\}$ . Elle est donc, par le point (1), de la forme  $R_0(\theta)$ , pour une certaine valeur de  $\theta$ . □

#### Proposition 18.128 ([609]).

Les rotations basées en  $(0, 0)$  forment un groupe abélien.

*Démonstration.* L'identité est une rotation par définition. En ce qui concerne l'inverse, si  $r = \sigma_1 \sigma_2$  alors  $r^{-1} = \sigma_2 \sigma_1$ . Nous commençons maintenant les choses pas tout à fait évidentes.

34. Définition 18.126.

- (1) **Composition** Soient des rotations  $r, r'$  centrées en  $(0, 0)$ . Soit également, une réflexion  $\sigma$  dont l'axe contient  $(0, 0)$ . Alors la proposition 18.125 nous donne l'existence de  $\sigma_1$  et  $\sigma_2$  tels que  $r = \sigma_1\sigma$  et  $r' = \sigma\sigma_2$ . Avec ça, la composition donne

$$rr' = \sigma_1\sigma\sigma\sigma_2 = \sigma_1\sigma_2, \quad (18.348)$$

qui est encore une rotation.

- (2) **Commutativité** Soient deux rotations  $r$  et  $r'$  ainsi que des décompositions  $r = \sigma_1\sigma$ ,  $r' = \sigma\sigma_2$ . Nous avons

$$rr' = \sigma_1\sigma_2 \quad (18.349a)$$

$$r'r = \sigma\sigma_2\sigma_1\sigma. \quad (18.349b)$$

Vu que  $t = \sigma_2\sigma_1$  est une rotation, nous pouvons encore appliquer la proposition 18.125 pour avoir  $t = \sigma_2\sigma_1 = \sigma\sigma_3$ . Ainsi,

$$r'r = \sigma\sigma\sigma_3\sigma = \sigma_3\sigma. \quad (18.350)$$

Mais aussi  $rr' = \sigma_1\sigma_2 = t^{-1} = \sigma_3\sigma$ . Nous avons donc bien  $rr' = r'r$ , et le groupe est commutatif.

□

NORMooOUDJooRfbDEX

### 18.129.

Jusqu'à présent nous avons parlé de rotations « affines ». Parmi elles, les rotations centrées en 0 (zéro, l'origine de  $E$  comme espace vectoriel) sont de particulière importance. Ce sont des applications linéaires, et même des isométries. Dans la suite, nous allons souvent dire simplement « rotation » pour dire « rotation centrée en 0 ».

Nous allons maintenant prendre un point de vue plus vectoriel, et allons noter les points de  $E$  avec des lettres comme  $x, y, u, v$  et non plus, avec des majuscules, comme lorsqu'on avait un point de vue affine. En même temps, nous allons noter les applications  $E \rightarrow E$  par des lettres comme  $A$  et ne plus écrire les parenthèses. Bref, nous écrivons  $Au$  au lieu de  $r(A)$ .

LEMooSYZYooWDFScw

### Lemme 18.130.

En dimension 2, les réflexions vectorielles (c'est-à-dire dont l'axe passe par 0) ont un déterminant qui vaut  $-1$ .

*Démonstration.* Soit une réflexion d'axe  $\ell$ . Prenons une base orthonormale de  $E$  constituée de  $e_1$  sur  $\ell$  et de  $e_2 \perp \ell$ . Alors  $\sigma_\ell(e_1) = e_1$  et  $\sigma_\ell(e_2) = -e_2$ . La formule du déterminant donne

$$\det(\sigma_\ell) = e_1^*(\sigma_\ell(e_1))e_2^*(\sigma_\ell(e_2)) - e_2^*(\sigma_\ell(e_1))e_1^*(\sigma_\ell(e_2)) = 1 \times (-1) - 0 \times 0 = -1. \quad (18.351)$$

Nous utilisons de façon cruciale le fait que le calcul du déterminant ne dépende pas de la base choisie, lemme 8.15.

□

PROPooTUJWooAjtEnQ

### Proposition 18.131.

Les rotations<sup>35</sup> sont

- (1) des applications linéaires orthogonales au sens de la définition 8.42,
- (2) des applications de déterminant 1,

*Démonstration.* Le fait qu'elles soient linéaires est la proposition 18.127.

Nous avons, pour tout  $u \in E$  l'égalité de la norme  $\|u\|$  et  $\|Au\|$  par le lemme 18.122 appliqué à  $Y = 0$ . En termes de produits scalaires nous avons alors  $\langle Au, Au \rangle = \langle u, u \rangle$ , et donc

$$\langle A^*Au, u \rangle = \|u\|^2. \quad (18.352)$$

35. Centrées en 0, nous ne le répéterons pas !

En particulier si  $\{e_i\}_{i=1,\dots,n}$  est une base orthonormée de  $E$  nous avons

$$(A^* A e_i)_i = \|e_i\|^2 = 1, \quad (18.353)$$

ce qui donne  $\|A^* A e_i\| \geq 1$ , avec égalité si et seulement si  $A^* A e_i = e_i$ . Ici nous avons utilisé le fait que  $\langle x, e_i \rangle = x_i$ , et le fait que pour tout  $i$  nous ayons  $\|x\| \geq |x_i|$ , avec égalité seulement si  $x$  est un multiple de  $e_i$ .

Par ailleurs l'inégalité de Cauchy-Schwarz 10.1 nous donne

$$\|u\|^2 = |\langle A^* A u, u \rangle| \leq \|A^* A u\| \|u\| \quad (18.354)$$

et donc

$$\|u\| \leq \|A^* A u\|. \quad (18.355)$$

Encore une fois, en appliquant cela à  $u = e_i$  nous trouvons  $1 \leq \|A^* A e_i\|$ . Puisque nous avons déjà l'inégalité dans l'autre sens,  $\|A^* A e_i\| = 1$ . Et le cas d'égalité est uniquement possible avec  $A^* A e_i = e_i$ .

Donc pour tout  $i$  de la base, nous avons  $A^* A e_i = e_i$ . Nous avons donc  $A^* A = \text{Id}$  et l'application  $A$  est orthogonale.

En ce qui concerne le déterminant, les réflexions sont de déterminant  $-1$  par le lemme 18.130, donc  $A = \sigma_1 \circ \sigma_2$  est de déterminant 1. Nous avons utilisé le fait que le déterminant était un morphisme : proposition 8.16(1).  $\square$

### Remarque 18.132.

Nous ne savons pas encore que les rotations forment tout le groupe  $\text{SO}(2)$  des endomorphismes orthogonaux de déterminant 1. Il faudra attendre le corolaire 18.137 pour le savoir.

LEM00MIJX00CjiQqP

### Lemme 18.133.

*L'application  $- \text{Id}$  est une rotation de  $\mathbb{R}^2$ .*

*Démonstration.* Soit une base orthonormée  $\{e_1, e_2\}$  de  $E$  et la rotation  $r = \sigma_1 \sigma_2$  où  $\sigma_i$  est la réflexion le long de l'axe  $\ell_i = \{te_i\}_{t \in \mathbb{R}}$ . Faut-il vous prouver que  $r = - \text{Id}$ ? La réflexion  $\sigma_2$  retourne la composante  $y$  d'un vecteur écrit dans la base  $\{e_1, e_2\}$  sans toucher à la composante  $x$ . La réflexion  $\sigma_1$  fait le contraire.  $\square$

## 18.10.7 Matrice des transformations orthogonales

Nous donnons maintenant une forme générale (trigonométrique) pour les matrices de  $\text{SO}(2)$ . Nous ne pouvons cependant pas invoquer les lemmes 18.134 ou 18.135 pour prétendre avoir une matrice des rotations, parce que nous n'avons pas encore prouvé que les rotations étaient des transformations orthogonales. Ce sera pour la proposition 18.138.

LEM00AJMA00XPSKtS

### Lemme 18.134.

*Si  $A \in \text{O}(2)$  alors il existe un unique  $(\theta, \epsilon) \in [0, 2\pi[ \times \{\pm 1\}$  tel que*

$$A = \begin{pmatrix} \cos(\theta) & -\epsilon \sin(\theta) \\ \sin(\theta) & \epsilon \cos(\theta) \end{pmatrix} \quad (18.356)$$

*Démonstration.* Soit une matrice  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  et imposons qu'elle soit dans  $\text{O}(2)$ . Le fait que  $A$  soit orthogonale impose

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (18.357)$$

Nous avons alors le système

$$\begin{cases} a^2 + c^2 = 1 & (18.358a) \\ b^2 + d^2 = 1 & (18.358b) \\ ab + cd = 0 & (18.358c) \end{cases}$$



La proposition 18.59 nous permet de déduire qu'il existe un unique  $\theta \in [0, 2\pi[$  tel que  $a = \cos(\theta)$ ,  $c = \sin(\theta)$ , ainsi que plusieurs  $\alpha \in \mathbb{R}$  tel que  $b = \cos(\alpha)$ ,  $d = \sin(\alpha)$ .

Note : si nous voulons  $\alpha \in [0, 2\pi[$ , alors il y a unicité. Ici nous ne nous attachons pas à cette contrainte ; nous savons qu'il en existe plusieurs, et nous allons en fixer un en fonction de  $\theta$ . Le  $\alpha$  ainsi fixé ne sera peut-être pas dans  $[0, 2\pi[$ , mais ce ne sera pas grave.

Les angles  $\theta$  et  $\alpha$  sont alors liés par la contrainte

$$\cos(\theta) \cos(\alpha) + \sin(\theta) \sin(\alpha) = 0. \quad (18.359)$$

Utilisant l'identité (18.22c) cela signifie que  $\cos(\theta - \alpha) = 0$ . Donc

$$\alpha \in \left\{ \theta + \frac{\pi}{2} + k\pi \right\}_{k \in \mathbb{Z}}. \quad (18.360)$$

Si  $k$  est pair, ça donne

$$\cos(\alpha) = -\sin(\theta) \quad (18.361a)$$

$$\sin(\alpha) = \cos(\theta) \quad (18.361b)$$

et alors

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (18.362)$$

Si au contraire  $k$  est impair, alors

$$\cos(\alpha) = \sin(\theta) \quad (18.363a)$$

$$\sin(\alpha) = -\cos(\theta), \quad (18.363b)$$

et

$$A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}. \quad (18.364)$$

Nous avons démontré qu'une matrice de  $O(2)$  était forcément d'une des deux formes (18.362) ou (18.364). Il est maintenant facile de vérifier que ces deux matrices sont effectivement dans  $O(2)$ .  $\square$

### Lemme 18.135.

Si  $g \in SO(2)$ , il existe un unique  $\theta \in [0, 2\pi[$  tel que la matrice de  $g$  dans la base canonique soit

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (18.365)$$

*Démonstration.* Puisque  $SO(2)$  est la partie de  $O(2)$  constituée des matrices de déterminant 1, nous pouvons reprendre la forme donnée par le lemme 18.134 et fixer  $\epsilon$  par la contrainte sur le déterminant. Nous avons, en utilisant la relation du lemme 18.4,

$$\det \begin{pmatrix} \cos(\theta) & -\epsilon \sin(\theta) \\ \sin(\theta) & \epsilon \cos(\theta) \end{pmatrix} = \epsilon. \quad (18.366)$$

Étant donné que  $\det(g) = 1$ , nous voyons que  $\epsilon = 1$ .  $\square$

### Corolaire 18.136 ([1]).

Nous avons une bijection

$$\begin{aligned} \psi: SO(2) &\rightarrow [0, 2\pi[ \\ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} &\mapsto \theta, \end{aligned} \quad (18.367)$$

et un isomorphisme de groupe

$$\begin{aligned} \varphi: SO(2) &\rightarrow U(1) = \{e^{i\theta}\}_{\theta \in \mathbb{R}} \\ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} &\mapsto e^{i\theta}. \end{aligned} \quad (18.368)$$



*Démonstration.* La première assertion est une paraphrase du lemme 18.135. Pour la seconde, il faut vérifier que c'est bien un morphisme et une bijection.

Pour le morphisme, ce sont les formules d'addition d'angle du lemme 18.14 qui jouent. En ce qui concerne la bijection...

(1) **Surjection** Comme  $e^{i\theta+2k\pi} = e^{i\theta}$ , tout élément de  $U(1)$  est exponentielle de  $i\theta$  pour un  $\theta \in [0, 2\pi[$ .

(2) **Injection** Remarquez que nous avons, pour  $\theta \in [0, 2\pi[$ ,  $\varphi(\psi^{-1}(\theta)) = e^{i\theta}$ .

Soient  $A, B \in SO(2)$  tels que  $\varphi(A) = \varphi(B)$ . Soient  $\theta_a, \theta_b \in [0, 2\pi[$  tels que  $A = \psi^{-1}(\theta_a)$  et  $B = \psi^{-1}(\theta_b)$ . L'hypothèse  $\varphi(A) = \varphi(B)$  donne

$$e^{i\theta_a} = e^{i\theta_b}, \quad (18.369)$$

et donc  $\theta_a = \theta_b$  par la proposition 18.62(1).

Nous en déduisons que  $A = B$ , et donc que  $\varphi$  est injective.

□

### 18.10.8 Rotations, $SO(2)$ et matrice de rotation

CORooVYUJooDbkIFY

**Corolaire 18.137** ([1]).

Le groupe des rotations centrées en  $(0, 0)$  est le groupe  $SO(2)$ .

*Démonstration.* Nous devons prouver deux choses :

- Toutes les rotations sont des éléments de  $SO(2)$ .
- Tous les éléments de  $SO(2)$  sont des rotations.

La proposition 18.127 nous indique que toute rotation de  $\mathbb{R}^2$  centrée en  $(0, 0)$  est de la forme  $R_0(\theta)$ , c'est-à-dire, a une matrice de la forme

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (18.370) \quad \text{EQooSJNBooMPIRZS}$$

Donc toute rotation est dans  $SO(2)$ .

D'autre part, le lemme 18.135 indique que tout élément de  $SO(2)$  a, dans la base canonique, une matrice de la forme (18.370). Le lemme 18.126 nous indique alors que c'est une rotation. □

PROPoOOTIVooZpvLnb

**Proposition 18.138.**

Si  $r$  est une rotation de  $\mathbb{R}^2$  centrée en  $(0, 0)$ , il existe un unique  $\theta \in [0, 2\pi[$  tel que  $r = R_0(\theta)$ .

*Démonstration.* Soit une rotation  $r$  autour de  $(0, 0)$ . Le corolaire 18.137 nous dit qu'il existe  $\theta \in \mathbb{R}$  tel que

$$r = R_0(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (18.371)$$

Nous avons identifié l'application linéaire à sa matrice. L'élément  $(\cos(\theta), \sin(\theta))$  est dans  $S^1$ , et il existe donc, par la proposition 18.59, un unique  $t \in [0, 2\pi[$  tel que  $\cos(t) = \cos(\theta)$  et  $\sin(t) = \sin(\theta)$ . Pour ce  $t$  nous avons alors

$$r = R_0(\theta) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}. \quad (18.372)$$

□

PROPoosUCooRYJcwo

**Proposition 18.139.**

Nous avons la formule suivante pour la composition :

$$R_0(\alpha) \circ R_0(\beta) = R_0(\alpha + \beta). \quad (18.373)$$

*Démonstration.* Par définition de  $R_0(\theta)$ , dans la base canonique de  $\mathbb{R}^2$ , la composition se calcule avec le produit suivant, en utilisant les formules du lemme 18.14 :

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} \quad (18.374a)$$

$$= \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\beta) \end{pmatrix} \quad (18.374b)$$

$$= \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}. \quad (18.374c)$$

Donc dans la base canonique, la matrice de  $R_0(\alpha)R_0(\beta)$  est celle de  $R_0(\alpha + \beta)$ .  $\square$

### 18.10.9 Rotation et application affine

Nous considérons à nouveau la définition 11.13 d'une application affine ainsi que sa décomposition en application linéaire et translation donnée par le lemme 11.14. Nous voyons maintenant comment ces propriétés se déclinent dans le cas d'une rotation non centrée en l'origine.

#### Exemple 18.140.

Soit  $A \in \mathbb{R}^2$  ainsi qu'une rotation  $f$  autour de  $A$ , c'est-à-dire une composition de deux symétries dont les axes se coupent en  $A$ <sup>36</sup>. Nous allons extraire de  $f$  la partie linéaire définie en 11.14.

Il existe des axes  $\ell_1$  et  $\ell_2$  tels que  $\ell_1 \cap \ell_2 = \{A\}$  et tels que

$$f = s_{\ell_1} \circ s_{\ell_2}. \quad (18.375)$$

En utilisant le lemme 18.116,

$$f = t_A \circ s_{t_A^{-1}(\ell_1)} \circ t_A^{-1} \circ t_A \circ s_{t_A^{-1}(\ell_2)} \circ t_A^{-1} = t_A \circ s_{t_A^{-1}(\ell_1)} \circ s_{t_A^{-1}(\ell_2)} \circ t_A^{-1}. \quad (18.376)$$

Comme  $A \in \ell_i$ , nous avons  $O = (0, 0) \in t_A^{-1}(\ell_i)$ . Donc les axes  $t_A^{-1}(\ell_1)$  et  $t_A^{-1}(\ell_2)$  se coupent en  $O$  et nous pouvons écrire

$$f = t_A \circ R \circ t_A^{-1} \quad (18.377)$$

où  $R$  est une rotation centrée en  $O$ ; donc une application linéaire par la proposition 18.127.

Nous avons

$$f(x) = (t_A \circ R \circ t_A^{-1})(x) \quad (18.378a)$$

$$= (t_A \circ R)(x - A) \quad (18.378b)$$

$$= t_A(R(x) - R(A)) \quad (18.378c)$$

$$= R(x) + t_{R(A)+A} \quad (18.378d)$$

$$= (t_{R(A)+A} \circ R)(x). \quad (18.378e)$$

Donc

$$f = t_{R(A)+A} \circ R. \quad (18.379)$$

Il est maintenant aisé de montrer que  $R$  est la partie linéaire de  $f$ . Pour tout  $M, x \in \mathbb{R}^2$  nous avons

$$f(M + x) = R(M + x) + R(A) + A \quad (18.380a)$$

$$= R(M) + R(x) + R(A) + A \quad (18.380b)$$

$$= R(x) + f(M). \quad (18.380c)$$

Donc nous confirmons la formule

$$f(M + x) = R(x) + f(M) \quad (18.381)$$

et  $R$  est la partie linéaire de  $f$ , voir la définition 11.14.  $\triangle$

<sup>36</sup>. Voir la définition 18.117.

LEMooUKEVooAEWvIM

**Lemme 18.141** ([1]).*Tout sous-groupe fini de  $\text{SO}(2)$  est cyclique*<sup>37</sup>.

*Démonstration.* Soit un sous-groupe fini  $G$  de  $\text{SO}(2)$ . Nous considérons les applications  $\psi: \text{SO}(2) \rightarrow [0, 2\pi[$  et  $\varphi: \text{SO}(2) \rightarrow U(1)$  données par le corolaire 18.136.

Vu que  $\psi(e) = 0$ , la partie  $\psi(G \setminus \{e\})$  est une partie finie de  $[0, 2\pi[$  qui ne contient pas 0. Nous en considérons le minimum  $\theta_0 \neq 0$ . Gardez en tête que ce minimum n'est pas nul :  $\theta_0 \neq 0$ .

Nous allons prouver que  $\psi^{-1}(\theta_0)$  génère le groupe  $G$ . Pour cela nous devons montrer que pour tout  $g \in G$ , il existe un entier  $k$  tel que  $g = \psi^{-1}(\theta_0)^k$ .

Soient  $g_1 \in \text{SO}(2)$  et  $\theta_1 = \psi(g_1)$ . Si il existe  $k \in \mathbb{N}$  tel que  $\theta_1 = k\theta_0$ , alors le résultat est prouvé. Supposons donc que  $\theta_1 = \lambda\theta_0$  avec  $\lambda$  non entier. Nous allons en déduire une contradiction.

Soit  $k$  la partie entière de  $\lambda$  (définition 1.582). Vu que  $\varphi$  est un morphisme de groupes, en utilisant la formule de la proposition 18.9(2) nous avons :

$$\varphi^{-1}(e^{i(\theta_1 - k\theta_0)}) = \varphi^{-1}(e^{i\theta_1} e^{-ik\theta_0}) \quad (18.382a)$$

$$= \varphi^{-1}(e^{i\theta_1}) \varphi^{-1}(e^{-ik\theta_0}) \quad (18.382b)$$

$$= \varphi^{-1}(e^{i\theta_1}) \varphi^{-1}(e^{t\theta_0})^k \quad (18.382c)$$

$$\in G. \quad (18.382d)$$

Nous avons donc  $\theta = \theta_1 - k\theta_0 \in \psi(G)$ . Mais en utilisant le lemme 1.584,

$$\theta - \theta_0 = \theta_1 - (k+1)\theta_0 < \theta_1 - \lambda\theta_0 = 0. \quad (18.383)$$

Cela signifie que  $\theta < \theta_0$ , et contredit la minimalité de  $\theta_0$ . □

### 18.10.10 Angle orienté

Avant d'aborder la classification des isométries, nous devons parler de l'angle entre deux droites. Si  $\ell_1$  et  $\ell_2$  sont deux droites, alors il est bien clair que deux angles peuvent prétendre être « l'angle entre  $\ell_1$  et  $\ell_2$  ». De plus chacun de ces deux angles sont doubles parce que si  $\alpha$  peut prétendre être l'angle entre  $\ell_1$  et  $\ell_2$ , alors  $-\alpha$  peut également y prétendre.

**Remarque 18.142.**

Nous ne parlons pas de l'angle entre  $\ell_1$  et  $\ell_2$  mais bien de l'angle *de*  $\ell_1$  *à*  $\ell_2$ . L'ordre des droites est important.

**18.143.**

Pour la suite,  $R_O(\alpha)$  est la rotation d'angle  $\alpha$  autour du point  $O$  tandis que  $R(\alpha)$  est la rotation d'angle  $\alpha$  autour de l'origine.

PROpooDWIMooQPkbw

**Proposition 18.144** ([610]).

*Si  $u$  et  $v$  sont des vecteurs unitaires*<sup>38</sup> *de  $\mathbb{R}^2$  alors il existe une unique rotation*<sup>39</sup> *$f$  telle que  $f(u) = v$ .*

*Démonstration.* C'est la proposition 18.123 appliquée à  $O = (0, 0)$ . □

**Remarque 18.145.**

Notons l'unicité. Nous ne faisons pas de différence entre  $R_\theta$  et  $R_{\theta+2\pi}$  et les autres  $R_{\theta+2k\pi}$ . En particulier si une rotation  $T$  est donnée, dire «  $T = R_\theta$  » ne définit pas un nombre  $\theta$  de façon univoque. Par contre ça définit une classe modulo  $2\pi$ , c'est-à-dire un élément  $\theta \in \mathbb{R}/2\pi$ .

Nous avons déjà défini le groupe  $\text{SO}(2)$  en la définition 8.46 et nous avons déterminé ses matrices dans  $\mathbb{R}^2$  en le lemme 18.135.

---

37. Définition 1.428.

38. De norme 1.

39. Définition 18.117.

La proposition 18.144 donne une application

$$T: S^1 \times S^1 \rightarrow \text{SO}(2). \quad (18.384)$$

Et nous avons une relation d'équivalence sur  $S^1 \times S^1$  donnée par  $(u, v) \sim (u', v')$  si et seulement si il existe  $g \in \text{SO}(2)$  telle que  $g(u) = u'$  et  $g(v) = v'$ .

DEFooVBKIoowIHvod

**Définition 18.146** (Angle orienté[610]).

Les classes de  $S^1 \times S^1$  pour cette relation d'équivalence sont les **angles orientés de vecteurs**. Nous notons  $[u, v]$  la classe de  $(u, v)$ .

PROFOoIWJQooGQJBWR

**Proposition 18.147.**

Nous avons  $T(u, v) = T(u', v')$  si et seulement si  $(u, v) \sim (u', v')$ .

*Démonstration.* En utilisant la commutativité du groupe  $\text{SO}(2)$  nous avons équivalence entre les affirmations suivantes :

- $(u, v) \sim (u', v')$
- $T(u, u') = T(v, v')$
- $T(u, u') \circ T(u', v) = T(v, v') \circ T(u', v)$
- $T(u, v) = T(u', v')$ .

□

**Proposition 18.148.**

Nous avons une bijection

$$S: \frac{S^1 \times S^1}{\sim} \rightarrow \text{SO}(2) \quad (18.385)$$

$$[u, v] \mapsto T(u, v).$$

*Démonstration.* En plusieurs points.

- (1) **S est bien définie** En effet si  $[u, v] = [z, t]$  alors  $T(u, v) = T(z, t)$ .
- (2) **Injectif** Si  $S[u, v] = S[z, t]$  alors  $T(u, v) = T(z, t)$ , qui implique  $(u, v) \sim (z, t)$  par la proposition 18.147.
- (3) **Surjectif** Nous avons  $R_\theta = T(u, R_\theta u)$ .

□

**Définition 18.149** (Somme d'angles orientés[610]).

Si  $[u, v]$  et  $[z, t]$  sont des angles orientés, nous définissons la somme par

$$[u, v] + [z, t] = S^{-1}(S[u, v] \circ S[z, t]). \quad (18.386)$$

LEMooWISVooYsStJp

**Lemme 18.150.**

Quelques propriétés des angles plats liées à la somme.

- (1)  $(S^1 \times S^1)/\sim$  est un groupe commutatif.

ITEMooBKTFooWbEvIU

- (2) Relations de Chasles :

$$[u, v] + [v, w] = [u, w]. \quad (18.387)$$

- (3)  $-[u, v] = [v, u]$ .

*Démonstration.* Pour la relation de Chasles, ça se base sur la propriété correspondante sur  $T$  :

$$[u, v] + [v, w] = S^{-1}(T(u, v) \circ T(v, w)) \quad (18.388a)$$

$$= S^{-1}(T(u, w)) \quad (18.388b)$$

$$= [u, w]. \quad (18.388c)$$

Pour l'inverse, la vérification est que

$$[u, v] + [v, u] = [u, u] = 0. \quad (18.389)$$

□

DEFooFLGNooCZUkHY

### Définition 18.151.

La **mesure** de l'angle orienté  $[u, v]$  est  $[\theta]_{2\pi}$  si  $T[u, v] = R_\theta$ .

Notons dans cette définition qu'écrire  $T[u, v] = R_\theta$  dans  $SO(2)$  ne définit pas  $\theta$ , mais seulement sa classe modulo  $2\pi$ . C'est pour cela que la mesure de l'angle orienté n'est également définie que modulo  $2\pi$ .

Pour la suite nous allons nous intéresser à des vecteurs qui ont, dans l'idée, un point de départ et un point d'arrivée. Si  $A, B \in \mathbb{R}^2$  nous notons

$$\overrightarrow{AB} = \frac{B - A}{\|B - A\|}. \quad (18.390)$$

C'est le vecteur unitaire dans la direction « de  $A$  vers  $B$  ».

THOooQDNKooTIVmmj

### Théorème 18.152 (Théorème de l'angle inscrit[611]).

Soit un cercle  $\Gamma$  de centre  $O$  et trois points distincts  $A, B, M \in \Gamma$ . Alors

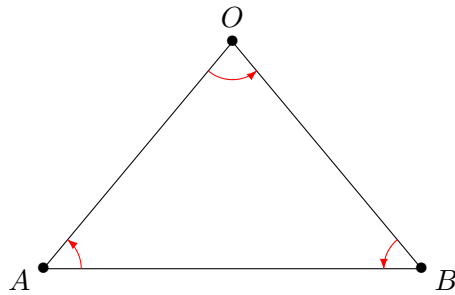
$$2(\overrightarrow{MA}, \overrightarrow{MB}) \in (\overrightarrow{OA}, \overrightarrow{OB})_{2\pi} \quad (18.391)$$

où l'indice  $2\pi$  indique la classe modulo  $2\pi$ .

*Démonstration.* Le triangle  $MOA$  est isocèle en  $O$ , donc les angles à la base sont égaux. Et de plus la somme des angles est dans  $[\pi]_{2\pi}$ . Bon, entre nous, nous savons que la somme des angles est exactement  $\pi$ , mais comme nous n'avons pas défini les angles autrement que modulo  $\pi$ , nous ne pouvons pas dire mieux. Donc

$$2(\overrightarrow{AB}, \overrightarrow{AO}) + (\overrightarrow{OB}, \overrightarrow{OA}) \in [\pi]_{2\pi}. \quad (18.392)$$

Il faut être sûr de l'orientation de tout cela. Le nombre  $(\overrightarrow{AB}, \overrightarrow{AO})$  est l'angle qui sert à amener  $\overrightarrow{AB}$  sur  $\overrightarrow{AO}$ . Vu que nous l'avons choisi dans le sens trigonométrique, il faut bien prendre les autres dans le sens trigonométrique, et utiliser  $(\overrightarrow{OA}, \overrightarrow{OB})$  et non  $(\overrightarrow{OB}, \overrightarrow{OA})$ .



De la même manière sur le triangle  $MOB$  nous écrivons

$$2(\overrightarrow{MB}, \overrightarrow{MO}) + (\overrightarrow{OM}, \overrightarrow{OB}) \in [\pi]_{2\pi}. \quad (18.393)$$

Nous faisons la différence entre les deux équations en remarquant que la différence de deux représentants de  $[\pi]_{2\pi}$  est un représentant de  $[0]_{2\pi}$  et en nous souvenant que  $-(\overrightarrow{MB}, \overrightarrow{MO}) = (\overrightarrow{MO}, \overrightarrow{MB})$  et les relations de Chasles du lemme 18.150(2) nous avons :

$$2(\overrightarrow{MA}, \overrightarrow{MB}) + (\overrightarrow{OB}, \overrightarrow{OA}) \in [0]_{2\pi}. \quad (18.394)$$

□

**18.153.**

Comment exprimer le fait qu'un angle orienté soit égal à  $\theta$  modulo  $\pi$  alors que les angles orientés sont des classes modulo  $2\pi$  ? Nous ne pouvons certainement pas écrire

$$(u, v) = [\theta]_\pi \quad (18.395)$$

parce que  $(u, v)$  est un élément de  $S^1 \times S^1$  alors que  $[\theta]_\pi$  est un ensemble de nombres. Nous pouvons écrire

$$[u, v] \subset [\theta]_\pi. \quad (18.396)$$

C'est cohérent parce que nous avons des deux côtés des ensembles de nombres. Les opérations permises sont l'égalité ou l'inclusion. L'égalité entre les deux ensembles n'est pas possible parce que la différence minimale entre deux éléments dans  $[u, v]$  est  $2\pi$  alors que celle dans  $[\theta]_\pi$  est  $\pi$ .

Si  $u$  et  $v$  forment un angle droit, nous avons

$$[u, v] = \left\{ \frac{\pi}{2} + 2k\pi \right\}_{k \in \mathbb{Z}}. \quad (18.397)$$

Et cela est bien un sous-ensemble de  $[\pi/2]_\pi$ .

Pour exprimer que deux angles orientés diffèrent de  $\pi$  nous devrions écrire :

$$[u, v] \subset [a, b]_\pi \quad (18.398)$$

où le membre de droite signifie la classe modulo  $\pi$  d'un représentant de  $[a, b]$ .

Nous allons cependant nous permettre d'écrire

$$[u, v] = [a, b]_\pi \quad (18.399)$$

voire carrément

$$(u, v) = (a, b)_\pi. \quad (18.400)$$

Cette dernière égalité devant être comprise comme voulant dire que l'angle pour passer de  $u$  à  $v$  est, soit le même que celui pour aller de  $a$  à  $b$ , soit ce dernier, plus  $\pi$ .

THOooUDUGooTJKDpO

**Théorème 18.154** ([611]).

Soient 4 points distincts du plan  $A, B, C, D$ . Ils sont alignés ou cocycliques<sup>40</sup> si et seulement si

$$(\overrightarrow{CA}, \overrightarrow{CB}) = (\overrightarrow{DA}, \overrightarrow{DB})_\pi. \quad (18.401)$$

Nous allons seulement démontrer l'implication directe.

*Démonstration.* Si les quatre points sont alignés nous avons  $[\overrightarrow{CA}, \overrightarrow{CB}] = [0]_{2\pi}$  et  $[\overrightarrow{DA}, \overrightarrow{DB}] = [0]_{2\pi}$ . En particulier nous avons

$$[\overrightarrow{CA}, \overrightarrow{CB}] = [\overrightarrow{DA}, \overrightarrow{DB}] \quad (18.402)$$

et a fortiori l'égalité modulo  $\pi$  au lieu de  $2\pi$ .

Nous nous relâchons en termes de notations. Si les quatre points sont cocycliques, nous pouvons utiliser le théorème de l'angle inscrit 18.152 dans les triangles  $ABC$  et  $ADB$  :

$$2(\overrightarrow{CA}, \overrightarrow{CB}) = (\overrightarrow{OA}, \overrightarrow{OB})_{2\pi} \quad (18.403a)$$

$$2(\overrightarrow{DA}, \overrightarrow{DB}) = (\overrightarrow{OA}, \overrightarrow{OB})_{2\pi}, \quad (18.403b)$$

ce qui donne  $2(\overrightarrow{CA}, \overrightarrow{CB}) = 2(\overrightarrow{DA}, \overrightarrow{DB})_{2\pi}$  et donc

$$(\overrightarrow{CA}, \overrightarrow{CB}) = (\overrightarrow{DA}, \overrightarrow{DB})_\pi. \quad (18.404)$$

Comme annoncé, nous ne faisons pas la preuve dans l'autre sens ; elle peut être trouvée dans [611]. □

---

40. C'est-à-dire sur un même cercle.

EXooOXAAooZMdDfP

**Exemple 18.155.**

À propos de groupe engendré et de générateur<sup>41</sup>. Soit  $G$  le groupe des rotations d'angle<sup>42</sup>  $k\pi/5$  (avec  $k$  entier). Ce groupe est constitué des « dixièmes de tour », puisque  $\frac{k\pi}{5} = \frac{2k\pi}{10}$ .

La rotation d'angle  $2\pi/5$  n'est pas génératrice parce qu'elle n'engendre que des « cinquièmes de tour » :  $4\pi/5$ ,  $6\pi/5$ ,  $8\pi/5$  et l'identité.

Par contre, la rotation d'angle  $\pi/5$  est génératrice.  $\triangle$

**18.10.11 Angles et nombres complexes**

SUBSECooKNUVooUBKaWm

Les nombres complexes peuvent être repérés par une norme et un angle, ce qui en fait un terrain propice à l'utilisation des angles orientés. Nous en ferons d'ailleurs usage dans  $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$  pour parler d'alignement, de cocyclicité et de birapport dans la proposition 23.88.

Soient deux éléments  $z_1, z_2 \in \mathbb{C}$ . Nous les écrivons sous la forme  $z_1 = r_1 e^{i\theta_1}$  et  $z_2 = r_2 e^{i\theta_2}$ ; remarquons que cela ne définit  $\theta_i$  qu'à  $2\pi$  près. Nous avons

$$[z_1, z_2] = [\theta_2 - \theta_1]_{2\pi}. \quad (18.405)$$

Soient maintenant  $a, b, c, d \in \mathbb{C}$ . Nous écrivons  $\vec{ab}$  le vecteur unitaire dans le sens « de  $a$  vers  $b$  », c'est-à-dire un multiple positif bien choisi du nombre  $b - a$ . Nous notons  $\theta_{ab}$  l'argument du nombre complexe  $b - a$ , et nous avons encore

$$[\vec{ab}, \vec{cd}] = [\theta_{ab} - \theta_{cd}]. \quad (18.406)$$

Avec toutes ces notations, ce qui est pas mal, c'est que les produits et quotients de nombres complexes se comportent très bien par rapport aux angles : l'argument de  $a/b$  est  $\theta_a - \theta_b$  et en particulier l'argument de

$$\frac{a - b}{c - d} \quad (18.407)$$

est dans la classe de l'angle orienté

$$[\vec{ba}, \vec{dc}]. \quad (18.408)$$

DEFooUPUooKAPFrh

**Définition 18.156** (Angle avec 3 points).

Soient trois points  $A, O, B \in \mathbb{R}^2$ . Voici comment nous définissons l'angle  $\widehat{AOB}$ ; de façon informelle, c'est l'angle de la rotation qui permet d'aller de  $A$  vers  $B$ .

- Nous nous mettons en l'origine :  $A' = A - O$  et  $B' = B - O$ .
- Nous normalisons :  $A'' = A'/\|A'\|$  et  $B'' = B'/\|B'\|$ .
- Soit  $f$ , l'unique rotation telle que  $f(A'') = B''$  (proposition 18.144).
- Soit  $\theta$  l'unique élément de  $[0, 2\pi[$  tel que la matrice de  $f$  dans la base canonique soit

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad (18.409)$$

par la proposition 18.138.

- L'angle  $\widehat{AOB}$  est ce nombre  $\theta$ .

Nous noterons également l'angle  $\widehat{AOB}$  par  $\text{angle}(A, O, B)$ .

Nous voyons que l'angle est toujours un nombre entre 0 et  $2\pi$ . Par abus de notation, nous admettrons de temps en temps, de parler d'angle en dehors de cet intervalle.

41. Définition 1.427 et 1.428

42. Voir la définition 18.151.

DEFooMAKWooGRutOk

**Définition 18.157** (Angle dans  $\mathbb{C}$ ).Si  $a, b, c \in \mathbb{C}$ , nous définissons l'angle  $\text{angle}(a, b, c) = \text{angle}(s(a), s(b), s(c))$  où

$$\begin{aligned} s: \mathbb{C} &\rightarrow \mathbb{R}^2 \\ x + iy &\mapsto (x, y), \end{aligned} \quad (18.410)$$

et le membre de droite étant défini par 18.156.

Avec cette définition, la multiplication dans  $\mathbb{C}$  par un élément de  $S^1$  revient à une rotation. C'est le lemme suivant.

LEMooUFOHooWdgIvO

**Lemme 18.158** ([1]).

Soit

$$\begin{aligned} s: \mathbb{C} &\rightarrow \mathbb{R}^2 \\ x + iy &\mapsto (x, y). \end{aligned} \quad (18.411)$$

Nous avons

$$s(e^{i\xi}z) = R_0(\xi)s(z) \quad (18.412)$$

pour tout  $\xi \in \mathbb{R}$  et  $z \in \mathbb{C}$ .

*Démonstration.* Nous écrivons  $z = re^{i\theta}$  comme l'autorise la proposition 18.63. En d'autres termes,  $z = r \cos(\theta) + ir \sin(\theta)$  et donc

$$s(z) = (r \cos(\theta), r \sin(\theta)). \quad (18.413)$$

D'autre part, la proposition 18.9 donne  $e^{i\xi}z = re^{i(\theta+\xi)}$ , et donc

$$s(e^{i\xi}z) = \begin{pmatrix} r \cos(\theta + \xi) \\ r \sin(\theta + \xi) \end{pmatrix}. \quad (18.414)$$

La conclusion est un calcul :

$$rR(\xi)(\cos(\theta), \sin(\theta)) = r \begin{pmatrix} \cos(\xi) & -\sin(\xi) \\ \sin(\xi) & \cos(\xi) \end{pmatrix} \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \quad \text{prop. 18.138} \quad (18.415a)$$

$$= r \begin{pmatrix} \cos(\xi) \cos(\theta) - \sin(\xi) \sin(\theta) \\ \sin(\xi) \cos(\theta) + \cos(\xi) \sin(\theta) \end{pmatrix} \quad (18.415b)$$

$$= r \begin{pmatrix} \cos(\xi + \theta) \\ \sin(\xi + \theta) \end{pmatrix} \quad \text{lemme 18.14} \quad (18.415c)$$

$$= s(e^{i\xi}z). \quad (18.415d)$$

□

LEMooXMYDooRrNblP

**Lemme 18.159.**Pour tout  $v \in S^1$ , l'application

$$\begin{aligned} s: [0, 2\pi[ &\rightarrow S^1 \\ \theta &\mapsto R(\theta)v \end{aligned} \quad (18.416)$$

est une bijection.

LEMooMFXWooGymSoF

**Lemme 18.160** ([1]).Si les points  $0$ ,  $R(\theta_1)v$  et  $R(\theta_2)v$  sont alignés, avec  $\theta_1 < \theta_2$  dans  $[0, 2\pi[$ , alors  $\theta_2 \in \{\theta_1, \theta_1 + \pi\}$ .

*Démonstration.* Supposons dans un premier temps que  $\|v\| = 1$ . Étant donné que  $R(\theta_2)v$  est sur la droite contenant  $0$  et  $R(\theta_1)v$ , il existe  $\lambda \in \mathbb{R}$  tel que

$$R(\theta_2)v = \lambda R(\theta_1)v. \quad (18.417)$$

Vu que  $R(\theta)$  est une isométrie, nous avons  $\lambda = \pm 1$ . Si  $\lambda = 1$ , alors nous avons  $R(\theta_1)v = R(\theta_2)v$  et donc  $\theta_1 = \theta_2$  par le lemme 18.159.

Si  $\lambda = -1$ , alors en remarquant que  $R(\pi)v = -v$ , nous avons  $R(\theta_1 + \pi)v = R(\theta_1)R(\pi)v = -R(\theta_1)v$ . Donc le lemme 18.159 nous dit que  $\theta_1 + \pi = \theta_2$ . □



DEFooEGKOooRPGOAs

**Lemme-Définition 18.161** (Angle entre deux droites[1]).

Soient deux droites  $\ell_1$  et  $\ell_2$  de  $\mathbb{R}^2$  sécantes au point  $A \in \mathbb{R}^2$ .

Il existe un unique  $\theta_0 \in [0, \pi[$  tel que<sup>43</sup>  $R_A(\theta_0)x \in \ell_2$  pour tout  $x \in \ell_1$ .

Cet angle  $\theta_0$  est l'**angle** de  $\ell_1$  à  $\ell_2$ .

*Démonstration.* Nous considérons le vecteur  $v$  tel que  $\ell_1 = \{A + tv\}_{t \in \mathbb{R}}$ , ainsi qu'un point arbitraire  $x_0 = A + \lambda v$  sur  $\ell_1$ . Nous allons montrer qu'il existe un unique  $\theta_\lambda \in ]0, \pi[$  tel que  $R_A(\theta_\lambda)x_0 \in \ell_2$ . Ensuite nous allons montrer que  $\theta_{\lambda_1} = \theta_{\lambda_2}$  pour tout  $\lambda_1, \lambda_2$  dans  $\mathbb{R}$ .

Nous considérons une application affine

$$\begin{aligned} f: \mathbb{R}^2 &\rightarrow \mathbb{R} \\ x &\mapsto u(x) + \alpha \end{aligned} \quad (18.418)$$

avec  $u: \mathbb{R}^2 \rightarrow \mathbb{R}$  linéaire et  $\alpha \in \mathbb{R}$  telle que  $\ell_2 = \ker(f)$  (proposition 12.153). Nous supposons que  $f(x_0) > 0$  pour fixer les idées.

Nous utilisons la formule (18.334) pour  $R_A$ ; pour tout  $\theta$  nous avons

$$R_A(\theta)(A + \lambda v) = (\tau_A \circ R(\theta) \circ \tau_A^{-1})(A + \lambda v) = \tau_A(R(\theta)(\lambda)v) = A + \lambda R(\theta)v. \quad (18.419)$$

Nous avons d'une part

$$f(x_0) = f(A + \lambda v) = u(A + \lambda v) + \alpha = u(A) + \alpha + \lambda u(v) = f(A) + \lambda u(v) = \lambda u(v) \quad (18.420)$$

parce que  $f(A) = 0$  du fait que  $A \in \ell_2$ .

De même, en tenant compte que  $R(\pi)v = -v$  nous avons

$$R_A(\pi)x_0 = u(-\lambda v + A) + \alpha = -\lambda u(v). \quad (18.421)$$

Donc  $f(R_A(\pi)x_0) = -f(x_0)$ .

Nous avons prouvé qu'en posant

$$\begin{aligned} s: [0, \pi] &\rightarrow \mathbb{R}^2 \\ \theta &\mapsto R_A(\theta)(A + \lambda v), \end{aligned} \quad (18.422)$$

nous avons  $s(0) = A + \lambda v$ ,  $s(\pi) = A - \lambda v$  et que  $(f \circ s)(0) > 0$  et  $(f \circ s)(\pi) < 0$ . Nous pouvons donc appliquer le théorème des valeurs intermédiaires 9.95 à la fonction continue  $f \circ s$ . Il existe  $\theta_0 \in ]0, \pi[$  tel que  $(f \circ s)(\theta_0) = 0$ .

Pour une telle valeur de  $\theta_0$ , nous avons  $s(\theta_0) \in \ker(f)$  et donc  $R_A(\theta_0)x_0 \in \ell_2$ .

- (1) **Unicité** Nous prouvons que si  $R_A(\theta_1)x_0$  et  $R_A(\theta_2)x_0$  sont sur  $\ell_2$ , alors  $\theta_1 = \theta_2$ . Supposons que ce les points  $A$ ,  $R_A(\theta_1)x_0$  et  $R_A(\theta_2)x_0$  sont alignés sur  $\ell_2$ . Alors les points  $0$ ,  $R(\theta_1)v$  et  $R(\theta_2)v$  sont alignés, et donc le lemme 18.160 nous assure que  $\theta_2 = \theta_1 + \pi$  ou que  $\theta_1 = \theta_2$ . La possibilité  $\theta_2 = \theta_1 + \pi$  est exclue parce que  $\theta_1$  et  $\theta_2$  sont dans  $]0, \pi[$ . Donc  $\theta_1 = \theta_2$ .

Nous avons vu qu'il existe une unique angle, que nous notons  $\theta_\lambda$  tel que  $R_A(\theta_\lambda)(A + \lambda v) \in \ell_2$ .

- (2) **Petite note au passage** La droite  $\ell_2$  est l'ensemble

$$\ell_2 = \{A + tR(\theta_\lambda)v\}_{t \in \mathbb{R}}. \quad (18.423)$$

- (3) **Ce  $\theta_\lambda$  fonctionne sur tout  $\ell_1$**  Tout élément de  $\ell_1$  peut être écrit sous la forme  $x_0 + \mu v$ . Nous prouvons à présent que si  $x_1 = x_0 + \mu v$ , alors  $R_A(\theta)x_1 \in \ell_2$ . Nous avons

$$R_A(\theta_\lambda)v = R_A(\theta_\lambda)(x_0 + \mu v) \quad (18.424a)$$

$$= R_A(\theta_\lambda)(A + (\lambda + \mu)v) \quad (18.424b)$$

$$= R(\theta_\lambda)((\lambda + \mu)v) + A \quad (18.424c)$$

$$= (\lambda + \mu)R(\theta_\lambda)v + A. \quad (18.424d)$$

Donc cela est dans  $\ell_2$ .

---

43. Pour rappel,  $R_A(\theta)$  est défini en 18.126.

□

PROPooKVSHooRODGWE

**Proposition 18.162** ([1]).*Les angles sont invariants sous les translations.**Plus précisément, si  $A, B, S, v \in \mathbb{R}^2$ , alors*

$$\widehat{T_v(A)T_v(S)T_v(B)} = \widehat{ASB} \quad (18.425)$$

*où  $T_v(X) = X + v$ .*

*Démonstration.* Nous notons  $X_v = X + v$ . Nous avons  $A'_v = A_v - S_v = (A + v) - (S + v) = A - S = A''$ . Donc les vecteurs  $A''$  et  $B''$  à partir desquels est calculé  $\widehat{ASB}$  sont les mêmes que les vecteurs  $A''_v$  et  $B''_v$  qui servent à calculer  $\widehat{T_v(A)T_v(S)T_v(B)}$ . □

PROPooYWKJooRjybUJ

**Proposition 18.163** ([1]).*Les angles sont invariants par rotations, c'est-à-dire que si  $A, B, S \in \mathbb{R}^2$  et si  $R_\theta$  est une rotation, alors*

$$\widehat{ASB} = \widehat{R_\theta(A)R_\theta(S)R_\theta(B)}. \quad (18.426)$$

*Démonstration.* Pour être plus concis, nous écrivons  $A_\theta$  for  $R_\theta(A)$  et de même pour  $B$  et  $S$ . Afin de calculer l'angle  $\widehat{R_\theta(A)R_\theta(S)R_\theta(B)}$ , nous définissons

$$\begin{cases} A'_\theta = A_\theta - S_\theta & (18.427a) \\ S'_\theta = 0 & (18.427b) \\ B'_\theta = B_\theta - S_\theta. & (18.427c) \end{cases}$$

et

$$\begin{cases} A''_\theta = \frac{A_\theta - S_\theta}{\|A_\theta - S_\theta\|} & (18.428a) \\ B''_\theta = \frac{B_\theta - S_\theta}{\|B_\theta - S_\theta\|}. & (18.428b) \end{cases}$$

Par définition, l'angle est le  $\alpha$  tel que  $R_\alpha(A''_\theta) = B''_\theta$ . Nous devons prouver que le même  $\alpha$  vérifie  $R_\alpha(A'') = B''$ .

Le fait que  $R_\theta$  soit une isométrie nous donne déjà

$$\|R_\theta(A) - R_\theta(B)\| = \|A - B\|. \quad (18.429)$$

Ensuite, la relation de définition de  $\alpha$  s'écrit

$$\frac{R_\alpha R_\theta(A) - R_\alpha R_\theta(S)}{\|A_\theta - S_\theta\|} = \frac{R_\theta(B) - R_\theta(S)}{\|B_\theta - S_\theta\|}. \quad (18.430)$$

Vu que  $R_\alpha$  et  $R_\theta$  commutent, nous avons

$$R_\theta \frac{R_\alpha(A) - R_\alpha(B)}{\|A - S\|} = R_\theta \frac{B - S}{\|B - S\|}, \quad (18.431)$$

et comme  $R_\theta$  est inversible, cela donne  $R_\alpha(A'') = B''$ . □

LEMooJLHGooQIpKIE

**Lemme 18.164** ([1]).*Soit  $A \in \mathbb{R}^2$  et une droite  $\ell_1$ . Soit  $\ell_2$  une droite passant par  $A$  et intersectant  $\ell_1$  en  $O$ . Alors*

$$\sigma_{\ell_1}(A) = R_O(-2\alpha)A \quad (18.432)$$

*où  $\alpha$  est l'angle de  $\ell_1$  à  $\ell_2$ .*

*Démonstration.* Nous allons utiliser des coordonnées autour de  $O$ . Il existe un vecteur  $v$  tel que

$$A = O + v \quad (18.433)$$

Par définition de l'angle  $\alpha$ <sup>44</sup>, la droite  $\ell_2$  s'obtient par rotation d'angle  $\alpha$  depuis la droite  $\ell_1$ . Donc le point

$$B = R_O(-\alpha)A \quad (18.434)$$

est sur  $\ell_1$ .

Nous allons prouver que le point

$$D = R_O(-2\alpha)A \quad (18.435)$$

est  $D = \sigma_{\ell_1}A$ .

Nous commençons par montrer que la droite  $(DA)$  est perpendiculaire à  $\ell_1$ , c'est-à-dire que

$$(D - A) \cdot (B - O) = 0. \quad (18.436)$$

En utilisant le fait que

$$R_O(\alpha)(O + X) = O + R(\alpha)X, \quad (18.437)$$

nous avons

$$D - A = R_O(-2\alpha)(O + v) - (O + v) = O + R(-2\alpha)v - O - v = R(-2\alpha)v - v \quad (18.438)$$

et de la même façon,

$$B - O = R(-\alpha)v. \quad (18.439)$$

Notons que tous les  $O$  se sont simplifiés et qu'il ne reste que des rotations usuelles. En utilisant le fait que  $R(\alpha)$  est une isométrie, nous pouvons alors calculer

$$(D - A) \cdot (B - O) = \langle R(-2\alpha)v - v, R(-\alpha)v \rangle \quad (18.440a)$$

$$= \langle R(-\alpha)v - R(\alpha)v, v \rangle. \quad (18.440b)$$

En utilisant la matrice de rotation du lemme 18.135 nous trouvons

$$(R(-\alpha) - R(\alpha))v = \begin{pmatrix} 2 \sin(\alpha)v_2 \\ -2 \sin(\alpha)v_1 \end{pmatrix} \quad (18.441)$$

et donc

$$\langle (R(-\alpha) - R(\alpha))v, v \rangle = 0. \quad (18.442)$$

Le point  $D$  est bien sûr la droite perpendiculaire à  $\ell_1$  et passant par  $A$ . Mais vu que  $D$  est obtenu à partir de  $A$  par une rotation, le point  $D$  est également sur le cercle de rayon  $\|OA\|$  et centré en  $O$ . Ce cercle possède exactement deux intersections avec cette droite. Le premier est  $A$  et le second est  $\sigma_{\ell_1}(A)$ . Comme  $D$  n'est pas  $A$ , nous avons  $D = \sigma_{\ell_1}(A)$ .  $\square$

### 18.10.12 Polygone convexe

DEFooBCSKooABKKtR

**Définition 18.165** ([612]).

Un *polygone* est une union de segments dans  $\mathbb{R}^2$  de la forme

$$[a_N, a_1] \cup \bigcup_{i=1}^N [a_i, a_{i+1}]. \quad (18.443)$$

La définition suivante dit ce qu'est un polygone convexe. À mon avis, le théorème de Jordan 27.216 doit permettre de donner une définition plus précise en disant que c'est la partie compacte qui est convexe.

---

44. Définition 18.161.

DEFooBCDPooXTpbFq

**Définition 18.166.**

Nous disons qu'un polygone est **convexe** si il est la frontière d'une partie convexe de  $\mathbb{R}^2$ .

PROFooFYRMooTqVDEm

**Proposition 18.167** ([613]).

Un polygone est convexe<sup>45</sup> si et seulement si il est contenu dans le demi-plan de chacun de ses côtés.

PROFooUPPTooZBFvPg

**Proposition 18.168.**

Les racines de l'unité dans  $\mathbb{C}$ , c'est-à-dire la partie

$$\{e^{2ik\pi/n}, k = 0, \dots, n-1\}, \quad (18.444)$$

forment un polygone régulier convexe<sup>46</sup>.

*Démonstration.* Nous posons  $a_k = e^{2ik\pi/n}$  et  $\xi = e^{2i\pi/n}$ , de telle sorte que  $a_{k+1} = \xi a_k$ . Le polygone que nous considérons est celui formé par les segments  $[a_i, a_{i+1}]$  avec  $i = 0, \dots, n-1$ . Nous savons que  $a_n = a_0$ ; donc nous pouvons parler seulement de segments de la forme  $[a_i, a_{i+1}]$  sans nous soucier du cas particulier  $[a_{n-1}, a_0]$ .

(1) **Longueurs** La longueur d'une arête est donnée par

$$|a_k - a_{k+1}| = |a_k - \xi a_k| = |a_k| |1 - \xi| = |1 - \xi| \quad (18.445)$$

parce que  $|a_k| = 1$ .

(2) **Angles** Nous avons un calcul utilisant pas mal de propriétés :

$$\text{angle}(a_l, a_{l+1}, a_{l+2}) = \text{angle}(\xi^l a_0, \xi^l a_1, \xi^l a_2) \quad (18.446a)$$

$$= \text{angle}(s(\xi^l a_0), s(\xi^l a_1), s(\xi^l a_2)) \quad \text{def. 18.157} \quad (18.446b)$$

$$= \text{angle}(R_0(\xi^l) s(a_0), R_0(\xi^l) s(a_1), R_0(\xi^l) s(a_2)) \quad \text{lem. 18.158} \quad (18.446c)$$

$$= \text{angle}(s(a_0), s(a_1), s(a_2)) \quad \text{prop. 18.163} \quad (18.446d)$$

$$= \text{angle}(a_0, a_1, a_2). \quad (18.446e)$$

(3) **Convexe** Voici un squelette de preuve<sup>47</sup>.

- Utiliser la caractérisation 18.167.
- Tous les sommets sont sur le cercle unité.
- Considérer  $d$ , la droite passant par deux point successifs  $e^{i\theta_1}$  et  $e^{i\theta_2}$ .
- Cette droite définit deux demi-plans.
- La proposition 12.167 devrait permettre de prouver que la partie du cercle entre  $\theta_1$  et  $\theta_2$  est dans un demi-plan, et le reste dans l'autre demi-plan.
- Il n'y a pas d'autres racines de l'unité entre  $\theta_1$  et  $\theta_2$ . Donc elles sont toutes dans le même demi-plan.
- Les demi-plans étant convexes (lemme 12.161), les arêtes sont dans le demi-plan.

□

45. Définition 18.166.

46. Polygone, définition 18.165, polygone convexe, définition 18.166. En ce qui concerne les angles, formellement c'est la définition 18.157.

47. Je ne n'ai pas vérifié. Écrivez-moi si vous voyez un problème.

### 18.10.13 Groupe diédral

#### 18.10.13.1 Définition et générateurs : vue géométrique

DEFooIWZGooAinSOh

##### Définition 18.169.

Le **groupe diédral**  $D_n$  ( $n \geq 3$ ) est le groupe des isométries de  $(\mathbb{C}, d)$  laissant invariant l'ensemble

$$\{e^{2ik\pi/n}, k = 0, \dots, n-1\} \quad (18.447)$$

des racines de l'unité.

##### 18.170.

La proposition 18.168 nous permet de dire que le groupe diédral est le groupe des isométries de  $\mathbb{R}^2$  laissant invariant un polygone régulier à  $n$  côtés. C'est un peu pour cela que nous n'avons défini  $D_n$  que pour  $n \geq 3$ ; et un peu aussi pour une raison technique qui arrivera quand on parlera des générateurs abstraits. Voir 18.179 et ce qui suit.

LEMooCUIVPooMZKnzo

##### Lemme 18.171.

Nous avons

$$D_n \subset O(2, \mathbb{R}). \quad (18.448)$$

*Démonstration.* Si  $f \in D_n$ , alors  $f(e^{2ik\pi/n})$  doit être l'un des  $e^{2ik'\pi/n}$ , et puisque  $f$  conserve les longueurs dans  $\mathbb{C}$ , nous devons avoir

$$1 = d(0, e^{2ik\pi/n}) = d(f(0), e^{2ik'\pi/n}). \quad (18.449)$$

Donc  $f(0)$  est à l'intersection de tous les cercles de rayon 1 centrés en les  $e^{2ik\pi/n}$ , ce qui montre que  $f(0) = 0$  (dès que  $n \geq 3$ ). Par conséquent notre étude du groupe diédral ne doit prendre en compte que les isométries vectorielles de  $\mathbb{R}^2$ . En d'autres termes

$$D_n \subset O(2, \mathbb{R}). \quad (18.450)$$

□

PROPooELOooVJtuZN

##### Proposition 18.172 ([614]).

Le groupe  $D_n$  contient un sous-groupe cyclique d'ordre 2 et un sous-groupe cyclique d'ordre  $n$ .

*Démonstration.* Nous notons  $s$  la conjugaison complexe<sup>48</sup>. C'est un élément d'ordre 2 qui est dans  $D_n$  parce que

$$s(e^{2ki\pi/n}) = e^{2(n-k)i\pi/n}. \quad (18.451) \quad \text{EqSUshknP}$$

De la même façon, la rotations d'angle  $2\pi/n$ , que l'on note  $r$ , agit sur les racines de l'unité et engendre le groupe d'ordre  $n$  des rotations d'angle  $2k\pi/n$ . □

##### Proposition 18.173 ([614]).

Si  $s$  est la conjugaison complexe et  $r$  la rotation d'angle  $2\pi/n$ , alors  $(sr)^2 = \text{Id}$ .

*Démonstration.* Si  $z^n = 1$ , alors

$$(sr sr)z = (sr s)(e^{2i\pi/n} z) = (sr)(e^{-2i\pi/n} \bar{z}) = s(\bar{z}) = z. \quad (18.452)$$

□

PropLDIPoZ

##### Proposition 18.174 ([614]).

Nous notons  $s$  la conjugaison complexe et  $r$  la rotation d'angle  $2\pi/n$ .

(1) Le groupe diédral  $D_n$  est engendré par  $s$  et  $r$ .

(2) Tous les éléments de  $D_n$  s'écrivent sous la forme  $r^m$  ou  $s \circ r^m$ .

ITEMooOEBHooULRmZk

48. C'est une réflexion; la réflexion d'axe  $\mathbb{R}$  dans  $\mathbb{C}$ .

*Démonstration.* Nous considérons les points  $A_0 = 1$  et  $A_k = e^{2ki\pi/n}$  avec  $k \in \{1, \dots, n-1\}$ . Par convention,  $A_n = A_0$ . L'action des éléments  $s$  et  $r$  sur ces points est

$$r(A_k) = A_{k+1} \quad (18.453a)$$

$$s(A_k) = A_{n-k}. \quad (18.453b)$$

Cette dernière est l'équation (18.451).

Soit  $f \in D_n$ . Étant donné que c'est une isométrie de  $\mathbb{R}^2$  avec un point fixe (le point 0),  $f$  est soit une rotation, soit une réflexion.

Supposons pour commencer qu'un des  $A_k$  soit fixé par  $f$ . Dans ce cas  $f$  a deux points fixes :  $O$  et  $A_k$  et est donc soit la réflexion d'axe  $(OA_k)$ , soit l'identité (lemme 18.114). L'identité étant engendrée par  $s$  et  $r$ , nous supposons que  $f$  est la réflexion d'axe  $(OA_k)$ . Dans ce cas, nous avons  $f = s \circ r^{n-2k}$ . En effet pour chaque  $l$  nous avons

$$(s \circ r^{n-2k})(A_l) = s(A_{l+n-2k}) = A_{n-l+n+2k} = A_{2k-l}. \quad (18.454)$$

Nous voyons que  $s \circ r^{n-2k}$  fixe  $O$  et  $A_k$ , et donc la droite  $(OA_k)$ . Elle est donc soit la réflexion d'axe  $(OA_k)$  soit l'identité. Mais elle n'est pas l'identité parce que elle ne fixe pas tous les  $A_l$ .

Nous passons à présent au cas où  $f$  ne fixe aucun des  $A_k$ .

(1) Supposons que  $f$  soit une rotation. Si  $f(A_k) = A_m$ , alors l'angle de la rotation est

$$\frac{2(m-k)\pi}{n}, \quad (18.455)$$

et donc  $f = r^{m-k}$ , qui est de la forme demandée.

(2) Supposons à présent que  $f$  soit une réflexion d'axe  $\Delta$ . Cette fois,  $\Delta$  ne passe par aucun des points  $A_k$ , par contre  $\Delta$  passe par 0. Nous commençons par montrer que  $\Delta$  doit être la médiatrice d'un des côtés  $[A_p, A_{p+1}]$  du polygone. Comme  $\Delta$  passe par  $O$  et n'est aucune des droites  $(OA_k)$ , cette droite passe par l'intérieur d'un des triangles  $OA_pA_{p+1}$  et intersecte donc le côté correspondant.

Notre tâche est de montrer que  $\Delta$  coupe  $[A_p, A_{p+1}]$  en son milieu. Dans ce cas,  $\Delta$  sera automatiquement perpendiculaire parce que le triangle  $OA_pA_{p+1}$  est isocèle en  $O$ . Nommons  $l$  la longueur des côtés du polygone  $P = \Delta \cap [A_p, A_{p+1}]$ ,  $x = d(A_p, P)$  et  $\delta = d(A_p, \Delta)$ . Vu que  $f$  est la symétrie d'axe  $\Delta$ , nous avons aussi  $d(f(A_p), \Delta) = \delta$  et  $d(A_p, f(A_p)) = 2\delta$ . D'autre part, par la définition de la distance,  $\delta < x$ . Si  $x < \frac{l}{2}$ , alors  $\delta < \frac{l}{2}$  et donc  $d(A_p, f(A_p)) < l$ . Or cela est impossible parce que le polygone ne possède aucun sommet à distance plus courte que  $l$  de  $A_p$ .

De la même manière si  $x > \frac{l}{2}$ , nous raisonnons avec  $A_{p+1}$  pour obtenir une contradiction. Nous en concluons que la seule possibilité est  $x = \frac{l}{2}$ , et donc  $f(A_p) = A_{p+1}$ . Montrons alors que  $f = s \circ r^{n-2p-1}$ . Il faut montrer que c'est une réflexion qui envoie  $A_p$  sur  $A_{p+1}$ . D'abord c'est une réflexion parce que

$$\det(sr^{n-2p-1}) = \det(s) \det(r^{n-2p-1}) = -1 \quad (18.456)$$

parce que  $\det(s) = -1$  alors que  $\det(r^k) = 1$  parce que  $r$  est une rotation dans  $SO(2)$ . Ensuite nous avons

$$s \circ r^{n-2p-1}(A_p) = s(A_{p+n-2p-1}) = s(A_{n-p-1}) = A_{n-(n-p-1)} = A_{p+1}. \quad (18.457)$$

Donc  $s \circ r^{n-2p-1}$  est bien une réflexion qui envoie  $A_p$  sur  $A_{p+1}$ .

□

CorWYITsWW

### Corolaire 18.175.

La liste des éléments de  $D_n$  est

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \quad (18.458)$$

et  $|D_n| = 2n$ .

*Démonstration.* Nous savons par la proposition 18.174 que tous les éléments de  $D_n$  s'écrivent sous la forme  $r^k$  ou  $sr^k$ . Puisque  $r$  est d'ordre  $n$ , il ne faut considérer que  $k \in \{1, \dots, n-1\}$ . Les éléments  $1, r, \dots, r^{n-1}$  sont tous différents, et sont (pour des raisons de déterminant) tous différents des  $sr^k$ . Les isométries  $sr^k$  sont toutes différentes entre elles pour essentiellement la même raison :

$$sr^k(A_p) = s(A_{p+k}) = A_{n-p+k} \quad (18.459)$$

donc si  $k \neq k'$ ,  $sr^k(A_p) \neq sr^{k'}(A_p)$ . La liste des éléments de  $D_n$  est donc

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \quad (18.460)$$

et donc  $|D_n| = 2n$ . □

EXooHNYyooUDsKnm

### Exemple 18.176.

Nous considérons le carré  $ABCD$  dans  $\mathbb{R}^2$  et nous cherchons les isométries de  $\mathbb{R}^2$  qui laissent le carré invariant. Nous nommons les points comme sur la figure 18.7. La symétrie d'axe vertical est nommée  $s$  et la rotation de 90 degrés est notée  $r$ .

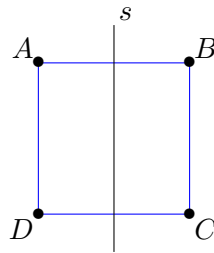


FIGURE 18.7: Le carré dont nous étudions le groupe diédral.

LabelFigIsomCarre

Il est facile de vérifier que toutes les symétries axiales peuvent être écrites sous la forme  $r^i s$ . De plus le groupe engendré par  $s$  agit sur le groupe engendré par  $r$  parce que

$$(srs^{-1})(A, B, C, D) = sr(B, A, D, C) = s(A, D, C, B) = (B, C, D, A), \quad (18.461)$$

c'est-à-dire  $srs^{-1} = r^{-1}$ . Nous sommes alors dans le cadre du corolaire 2.50 et nous pouvons écrire que

$$D_4 = \text{gr}(r) \rtimes_{\sigma} \text{gr}(s). \quad (18.462)$$

△

### 18.10.13.2 Table de multiplication

La proposition 18.174 nous indique que tous les éléments de  $D_n$  s'écrivent sous la forme  $s^{\epsilon} \circ r^m$  avec  $\epsilon \in \{0, 1\}$ . Nous allons maintenant écrire la table de multiplication pour de telles transformations de  $\mathbb{C}$ .

LEMooBNJFooAbhsUa

#### Lemme 18.177.

Si  $R$  est une rotation autour de 0 (dans  $\mathbb{C}$ ), et si  $s$  est la conjugaison complexe, alors

$$Rs = sR^{-1} \quad (18.463)$$

*Démonstration.* Il s'agit seulement d'un calcul en écrivant  $R$  comme la multiplication par  $e^{i\alpha}$ . Nous avons

$$(Rs)z = e^{i\alpha} \bar{z} = s(e^{-i\alpha} z) = sR^{-1}z. \quad (18.464)$$

□

PROPooPYDLooLgiUjk

#### Proposition 18.178.

Si  $\epsilon_1, \epsilon_2 \in \{0, 1\}$  et si  $k, l \in \mathbb{Z}$  nous avons

$$(s^{\epsilon_1} r^k)(s^{\epsilon_2} r^l) = s^{\epsilon_1 + \epsilon_2} r^{l + (-1)^{\epsilon_1} k}. \quad (18.465)$$

*Démonstration.* Si  $\epsilon_2 = 1$  alors nous utilisons le lemme 18.177 pour trouver

$$(s^{\epsilon_1} r^k)(s^{\epsilon_2} r^l) = s^{\epsilon_1} (r^k s^{\epsilon_2}) r^l = s^{\epsilon_1} s^{\epsilon_2} r^{-k} r^l. \quad (18.466)$$

La proposition est déjà prouvée dans ce cas.

Passons à  $\epsilon_2 = 0$ . Dans ce cas nous avons

$$(s^{\epsilon_1} r^k)(s^{\epsilon_2} r^l) = s^{\epsilon_1} r^{k+l}, \quad (18.467)$$

et c'est bon.  $\square$

### 18.10.13.3 Générateurs : vue abstraite

NORMMooCCUEooRRENed

#### 18.179.

Nous allons montrer que  $D_n$  peut être décrit de façon abstraite en ne parlant que de ses générateurs. Nous considérons un groupe  $G$  engendré par des éléments  $a$  et  $b$  tels que

- (1)  $a$  est d'ordre 2,
- (2)  $b$  est d'ordre  $n$  avec  $n \geq 3$ ,
- (3)  $abab = e$ .

Nous allons prouver que ce groupe doit avoir la même liste d'éléments que celle du corolaire 18.175.

**Proposition 18.180** ([614]).

*Le groupe  $G$  n'est pas abélien.*

*Démonstration.* Nous savons que  $abab = e$ , donc  $abab^{-1} = b^{-2}$ , mais  $b^{-2} \neq e$  parce que  $b$  est d'ordre  $n > 2$ . Donc  $abab^{-1} \neq e$ . En manipulant un peu :

$$e \neq abab^{-1} = (ab)(ba^{-1})^{-1} = (ab)(ba)^{-1} \quad (18.468)$$

parce que  $a^{-1} = a$ . Donc  $ab \neq ba$ .  $\square$

LemKKXdqdL

**Lemme 18.181** ([614]).

*Pour tout  $k$  entre 1 et  $n - 1$  nous avons*

$$Ad(a)b^k = ab^k a^{-1} = ab^k a = b^{-k}. \quad (18.469)$$

*Démonstration.* Nous faisons la démonstration par récurrence. D'abord pour  $k = 1$ , nous devons avoir  $aba = b^{-1}$ , ce qui est correct parce que par construction de  $G$  nous avons  $abab = e$ . Ensuite nous supposons que le lemme tient pour  $k$  et nous regardons ce qu'il se passe avec  $k + 1$  :

$$ab^{k+1}a = ab^kba = \underbrace{ab^k a}_{b^{-k}} \underbrace{aba}_{b^{-1}} = b^{-k}b^{-1} = b^{-(k+1)}. \quad (18.470)$$

$\square$

PROPoovQARooWuKHMZ

**Proposition 18.182.**

*L'élément  $a$  n'est pas une puissance de  $b$ .*

*Démonstration.* Supposons le contraire :  $a = b^k$ . Dans ce cas nous aurions

$$e = (ab)(ab) = b^{k+1}b^{k+1} = b^{2k+2} = b^{2k}b^2 = a^2b^2 = b^2, \quad (18.471)$$

ce qui signifierait que  $b$  est d'ordre 2, ce qui est exclu par construction.  $\square$

PROPoovEPVGooQjHRJp

**Proposition 18.183** ([614]).

*La liste des éléments de  $G$  est donnée par*

$$G = \{1, b, \dots, b^{n-1}, a, ab, \dots, ab^{n-1}\} = \{a^\epsilon b^k\}_{\substack{\epsilon=0,1 \\ k=0,\dots,n-1}} \quad (18.472)$$

*Les éléments de ces listes sont distincts.*



*Démonstration.* Étant donné que  $a$  n'est pas une puissance de  $b$ , les éléments  $1, a, b, \dots, b^{n-1}$  sont distincts. De plus si  $k$  et  $m = k + p$  sont deux éléments distincts de  $\{1, \dots, n-1\}$ , nous avons  $ab^k \neq ab^m$  parce que si  $ab^k = ab^{k+p}$ , alors  $a = ab^p$  avec  $p < n$ , ce qui est impossible. Pour la même raison,  $ab^k \neq e$ , et  $ab^k \neq b^m$ .

Au final les éléments  $1, a, b, \dots, b^{n-1}, ab, \dots, ab^{n-1}$  sont tous différents. Nous devons encore voir qu'il n'y en a pas d'autres.

Par définition le groupe  $G$  est engendré par  $a$  et  $b$ , donc tout élément  $x \in G$  s'écrit  $x = a^{m_1} b^{k_1} \dots a^{m_r} b^{k_r}$  pour un certain  $r$  et avec pour tout  $i$ ,  $k_i \in \{1, \dots, n-1\}$  (sauf  $k_r$  qui peut être égal à zéro) et  $m_i = 1$ , sauf  $m_1$  qui peut être égal à zéro. Donc

$$x = a^m b^{k_1} a b^{k_2} a \dots b^{k_{r-1}} a b^{k_r} \quad (18.473)$$

où  $m$  et  $k_r$  peuvent éventuellement être zéro. En utilisant le lemme 18.181 sous la forme  $b^{k_i} a = ab^{-k_i}$ , quitte à changer les valeurs des exposants, nous pouvons passer tous les  $a$  à gauche et tous les  $b$  à droite pour finir sous la forme  $x = a^k b^m$ .

Donc non, il n'existe pas d'autres éléments dans  $G$  que ceux déjà listés.  $\square$

LemmoNFRloPwUikh

**Lemme 18.184** ([1]).

Tout élément de  $G$  s'écrit de façon unique sous la forme  $a^\epsilon b^k$  ou  $b^k a^\epsilon$  avec  $\epsilon \in \{0, 1\}$  et  $k \in \{0, \dots, n-1\}$ .

*Démonstration.* Nous commençons par la forme  $a^\epsilon b^k$ . L'existence est la proposition 18.183. Pour l'unicité nous supposons  $a^\epsilon b^k = a^\sigma b^l$  et nous décomposons en 4 cas distincts.

- (1)  $\epsilon = 0, \sigma = 0$  Alors  $b^k = b^l$ . Mais  $b$  étant d'ordre  $n$  et  $k, l$  étant égaux au maximum à  $n-1$ , cette égalité implique  $k = l$ .
- (2)  $\epsilon = 0, \sigma = 1$  Alors  $b^k = ab^l$ , ce qui donne  $a = b^{k-l}$ , ce qui est interdit par la proposition 18.182.
- (3)  $\epsilon = 1, \sigma = 0$  Même problème que ci-dessus.
- (4)  $\epsilon = 1, \sigma = 1$  Encore une fois  $b^k = b^l$  implique  $k = l$ .

En ce qui concerne la forme  $b^k a^\epsilon$ , l'existence est à montrer. Soit l'élément  $g = a^\epsilon b^k$ ; cherchons à le mettre sous la forme  $b^l a^\sigma$ . Si  $\epsilon = 0$  c'est évident. Sinon  $\epsilon = 1$  et nous avons par le lemme 18.181

$$ab^k = b^{-k} a^{-1} = b^{-k} b^n a = b^{-k} a. \quad (18.474)$$

En ce qui concerne l'unicité, nous distinguons 4 cas pour  $b^k a^\epsilon = b^l a^\sigma$ . Comme précédemment ils se traitent exactement comme précédemment.  $\square$

THOooYITHooTNTBuG

**Théorème 18.185.**

Les groupes  $G$  et  $D_n$  sont isomorphes.

*Démonstration.* Nous utilisons l'application

$$\begin{aligned} \psi: G &\rightarrow D_n \\ a^k b^m &\mapsto s^k r^m. \end{aligned} \quad (18.475)$$

C'est évidemment bien défini et bijectif, mais c'est également un morphisme parce que si nous calculons  $\psi$  sur un produit, nous devons comparer

$$\psi(a^{k_1} b^{m_1} a^{k_2} b^{m_2}) \quad (18.476)$$

EoBUlPih

avec

$$\psi(a^{k_1} b^{m_1}) \psi(a^{k_2} b^{m_2}) = s^{k_1} r^{m_1} s^{k_2} r^{m_2}. \quad (18.477)$$

EoIVELph

Vu que  $D_n$  et  $G$  ont les mêmes propriétés qui permettent de permuter  $a$  et  $b$  ou  $s$  et  $r$ , l'expression à l'intérieur du  $\psi$  dans (18.476) se simplifie en  $a^k b^m$  avec les même  $k$  et  $m$  que l'expression à droite dans (18.477) et se simplifie en  $s^k r^m$ .  $\square$

**Corolaire 18.186.**

Toutes les propriétés démontrées pour  $G$  sont vraies pour  $D_n$ . En particulier, avec quelques redites :

- (1) Le groupe  $D_n$  peut être défini comme étant le groupe engendré par un élément  $s$  d'ordre 2 et un élément  $r$  d'ordre  $n - 1$  assujettis à la relation  $sr sr = e$ .
- (2) Le groupe  $D_n$  n'est pas abélien.
- (3) Pour tout  $k \in \{1, \dots, n - 1\}$  nous avons  $sr^k s = r^{-k}$ .
- (4) L'élément  $s$  ne peut pas être obtenu comme une puissance de  $r$ .
- (5) La liste des éléments de  $D_n$  est

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \quad (18.478)$$

- (6) Le groupe diédral  $D_n$  est d'ordre  $2n$ .

**Proposition 18.187.**

En posant  $C_n = \{r^k\}_{k=0, \dots, n-1}$  et  $C_2 = \{a^\epsilon\}_{\epsilon=0,1}$ , nous pouvons exprimer  $D_n$  comme le produit semi-direct

$$D_n = C_n \times_\rho C_2 \quad (18.479)$$

où  $\rho$  désigne l'action adjointe.

*Démonstration.* L'isomorphisme est :

$$\begin{aligned} \psi: C_n \times_\rho C_2 &\rightarrow D_n \\ (b^k, a^\epsilon) &\mapsto b^k a^\epsilon. \end{aligned} \quad (18.480)$$

- (1) **Action adjointe** L'application  $\rho_{a^\epsilon} = \mathbf{Ad}(a^\epsilon)$  est toujours un morphisme. Comme  $a^\epsilon$  est, soit  $e$ , soit  $a$ , nous allons nous restreindre à  $a$  et oublier l'exposant  $\epsilon$ . Il faut montrer que  $\mathbf{Ad}(a) \in \text{Aut}(C_n)$ . En utilisant le lemme 18.181,

$$\mathbf{Ad}(a)b^k = ab^k a^{-1} = b^{-k} = b^{n-k}. \quad (18.481)$$

L'application  $\mathbf{Ad}(a): C_n \rightarrow C_n$  est donc bijective et homomorphique. Ergo isomorphisme.

- (2) **Injectif** Si  $\psi(b^k, a^\epsilon) = \psi(b^l, a^\sigma)$ , alors par unicité du lemme 18.184 nous avons  $k = l$  et  $\epsilon = \sigma$ .
- (3) **Surjectif** Par la partie « existence » du lemme 18.184.
- (4) **Morphisme** Lorsqu'on prend deux sous-groupes d'un même groupe (ici le groupe des isométries de  $\mathbb{R}^2$ ), et que l'on tente de faire un produit semi-direct en utilisant l'action adjointe, nous avons toujours un morphisme. Dans notre cas, le calcul est :

$$\psi((b^k, a^\epsilon)(b^l, a^\sigma)) = b^k \rho_{a^\epsilon}(b^l) a^{\epsilon+\sigma} = b^k a^\epsilon b^l a^{-\epsilon} a^{\epsilon+\sigma} = b^k a^\epsilon b^l a^\sigma = \psi(b^k, a^\epsilon) \psi(b^l, a^\sigma). \quad (18.482)$$

□

**18.10.13.4 Classes de conjugaison**

subsubsecZQnBcgo

Pour les classes de conjugaison du groupe diédral nous suivons [615].

D'abord pour des raisons de déterminants<sup>49</sup>, les classes des éléments de la forme  $r^k$  et de la forme  $sr^k$  ne se mélangent pas. Nous notons  $C(x)$  la classe de conjugaison de  $x$ , et  $y \cdot x = yxy^{-1}$ .

Les relations que nous allons utiliser sont

$$sr^k s = r^{-k} \quad (18.483a)$$

$$rs = sr^{-1} = sr^{n-1}. \quad (18.483b)$$

49. Vous notez qu'ici nous utilisons un argument qui utilise la définition de  $D_n$  comme isométries de  $\mathbb{R}^2$ . Si nous avions voulu à tout prix nous limiter à la définition « abstraite » en termes de générateurs, il aurait fallu trouver autre chose.

La classe de conjugaison qui ne rate jamais est bien entendu  $C(1) = 1$ . Nous commençons les vraies festivités avec  $C(r^m)$ . D'abord  $r^k \cdot r^m = r^m$ , ensuite

$$(sr^k) \cdot r^m = sr^k r^m r^{-k} s^{-1} = sr^m s^{-1} = r^{-m}. \quad (18.484)$$

Donc

$$C(r^m) = \{r^m, r^{-m}\}. \quad (18.485)$$

À ce niveau il faut faire deux remarques. D'abord si  $m > \frac{n}{2}$ , alors  $C(r^m)$  est la classe de  $C^{n-m}$  avec  $n - m < \frac{n}{2}$ . Donc les classes que nous avons trouvées sont uniquement à lister avec  $m < \frac{n}{2}$ . Ensuite si  $m = \frac{n}{2}$  alors  $r^m = r^{-m}$  et la classe est un singleton. Cela n'arrive que si  $n$  est pair.

Nous passons ensuite à  $C(s)$ . Nous avons

$$r^k \cdot s = r^k sr^{-k} = ssr^k sr^{-k} = sr^{-k} r^{-k} = sr^{n-2k}, \quad (18.486)$$

et

$$(sr^k) \cdot s = \underbrace{sr^k s}_{r^{-k}} r^{-k} s^{-1} = r^{-2k} s = r^{n-2k} s = sr^{(n-1)(n-2k)} = sr^{n^2-2kn-n+2k} = sr^{2k}. \quad (18.487)$$

donc

$$C(s) = \{sr^{n-2k}, sr^{2k}\}_{k=0,\dots,n-1}. \quad (18.488)$$

Ici aussi l'écriture n'est pas optimale : peut-être que pour certains  $k$  il y a des doublons. Nous reportons l'écriture exacte à la discussion plus bas qui distinguera  $n$  pair de  $n$  impair. Notons juste que si  $n$  est pair, l'élément  $sr$  n'est pas dans la classe  $C(s)$ .

Nous en faisons donc à présent le calcul en gardant en tête le fait qu'il n'a de sens que si  $n$  est pair. D'abord

$$s \cdot (sr) = ssrs = rs = sr^{n-1}. \quad (18.489)$$

Ensuite

$$(sr^k) \cdot (sr) = sr^k sr r^{-k} s = r^{-2k+1} s = sr^{2k-1}. \quad (18.490)$$

Avec  $k = \frac{n}{2}$ , cela rend  $s \cdot (sr)$ , donc pas besoin de le recopier. Nous avons

$$C(sr) = \{sr^{2k-1}\}_{k=1,\dots,n-1}. \quad (18.491)$$

### 18.10.13.5 Le compte pour $n$ pair

SubsubsecROVmHuM

Si  $n$  est pair, nous avons les classes

$$C(1) = \{1\} \quad 1 \text{ élément} \quad (18.492a)$$

$$C(r^m) = \{r^m, r^{-m}\} = \{r^m, r^{n-m}\} \quad \text{pour } 0 < m < \frac{n}{2} \quad \frac{n}{2} - 1 \text{ fois } 2 \text{ éléments} \quad (18.492b)$$

$$C(r^{n/2}) = \{r^{n/2}\} \quad 1 \text{ élément} \quad (18.492c)$$

$$C(s) = \{sr^{2k}\}_{k=0,\dots,\frac{n}{2}-1} \quad \frac{n}{2} \text{ éléments} \quad (18.492d)$$

$$C(sr) = \{sr^{2k+1}\}_{k=0,\dots,\frac{n}{2}-1} \quad \frac{n}{2} \text{ éléments.} \quad (18.492e)$$

Au total nous avons bien listé  $2n$  éléments comme il se doit, dans  $\frac{n}{2} + 3$  classes différentes.

### 18.10.13.6 Le compte pour $n$ impair

GJIzDEP

Si  $n$  est impair, nous avons les classes

$$C(1) = \{1\} \quad 1 \text{ élément} \quad (18.493a)$$

$$C(r^m) = \{r^m, r^{-m}\} = \{r^m, r^{n-m}\} \quad \text{pour } 0 < m < \frac{n-1}{2} \quad \frac{n-1}{2} \text{ fois } 2 \text{ éléments} \quad (18.493b)$$

$$C(s) = \{sr^k\}_{k=0,\dots,n-1} \quad n \text{ éléments} \quad (18.493c)$$

Au total nous avons bien listé  $2n$  éléments comme il se doit, dans  $\frac{n+3}{2}$  classes différentes.

### 18.10.14 Applications : du dénombrement

#### 18.10.14.1 Le jeu de la roulette

pTqJLY

Soit une roulette à  $n$  secteurs que nous voulons colorier en  $q$  couleurs [616]. Nous voulons savoir le nombre de possibilités à rotations près. Soit d'abord  $E$  l'ensemble des coloriages possibles sans contrainte ; il y a naturellement  $q^n$  possibilités. Sur l'ensemble  $E$ , le groupe cyclique  $G$  des rotations d'angle  $2\pi/n$  agit. Deux coloriages étant identiques si ils sont reliés par une rotation, la réponse à notre problème est donnée par le nombre d'orbites de l'action de  $G$  sur  $E$  qui sera donnée par la formule du théorème de Burnside 2.41.

Nous devons calculer  $\text{Card}(\text{Fix}(g))$  pour tout  $g \in G$ . Soit  $g$ , un élément d'ordre  $d$  dans  $G$ . Si  $g$  agit sur la roulette, chaque secteur a une orbite contenant  $d$  éléments. Autrement dit,  $g$  divise la roulette en  $n/d$  secteurs. Un élément de  $E$  appartenant à  $\text{Fix}(g)$  doit colorier ces  $n/d$  secteurs de façon uniforme ; il y a  $q^{n/d}$  possibilités.

Il reste à déterminer le nombre d'éléments d'ordre  $d$  dans  $G$ . Un élément de  $G$  est donné par un nombre complexe de la forme  $e^{2ik\pi/n}$ . Les éléments d'ordre  $d$  sont les racines primitives<sup>50</sup>  $d$ -ièmes de l'unité. Nous savons que –par définition– il y a  $\varphi(d)$  telles racines primitives de l'unité. Bref il y a  $\varphi(d)$  éléments d'ordre  $d$  dans  $G$ .

La formule de Burnside nous donne maintenant le nombre d'orbites :

$$\frac{1}{n} \sum_{d|n} \varphi(d) q^{n/d}. \quad (18.494)$$

Cela est le nombre de coloriage possibles de la roulette à  $n$  secteurs avec  $q$  couleurs.

#### 18.10.14.2 L'affaire du collier

siOQIG

Nous avons maintenant des perles de  $q$  couleurs différentes et nous voulons en faire un collier à  $n$  perles. Cette fois non seulement les rotations donnent des colliers équivalents, mais en outre les symétries axiales (il est possible de retourner un collier, mais pas une roulette). Le groupe agissant sur  $E$  est maintenant le groupe diédral<sup>51</sup>  $D_n$  conservant un polygone à  $n$  sommets.

Nous devons séparer le cas  $n$  impair, du cas  $n$  pair.

Si  $n$  est impair, alors les axes de symétries passent par un sommet et par le milieu du côté opposé. Le groupe  $D_n$  contient  $n$  symétries axiales. Nous avons donc maintenant

$$|G| = 2n. \quad (18.495)$$

Nous écrivons la formule de Burnside

$$\text{Card}(\Omega) = \frac{1}{2n} \sum_{g \in G} \text{Card}(\text{Fix}(g)). \quad (18.496)$$

Si  $g$  est une rotation, le travail est déjà fait. Si  $g$  est une symétrie, nous avons le choix de la couleur du sommet par lequel passe l'axe et le choix de la couleur des  $(n-1)/2$  paires de sommets. Cela fait

$$qq^{(n-1)/2} = q^{\frac{n+1}{2}} \quad (18.497)$$

possibilités. Nous avons donc

$$\text{Card}(\Omega) = \frac{1}{2n} \left( \sum_{d|n} q^{n/d} \varphi(d) + nq^{\frac{n+1}{2}} \right). \quad (18.498)$$

Si  $n$  est pair, les choses se compliquent un tout petit peu. En plus de symétries axiales passant par les milieux de deux côtés opposés, il y a les axes passant par deux sommets opposés. Pour

50. Une racine non primitive 8ième de l'unité est par exemple  $i$ . Certes  $i^8 = 1$ , mais  $i^4 = 1$  aussi. Le nombre  $i$  est d'ordre 4.

51. Définition 18.169.

colorier un collier en tenant compte d'une telle symétrie, nous pouvons choisir la couleur des deux perles par lesquelles passe l'axe ainsi que la couleur des  $(n-2)/2$  paires de perles. Cela fait en tout

$$q^2 q^{\frac{n-2}{2}} = q^{\frac{n+2}{2}}. \quad (18.499)$$

Le groupe  $G$  contient  $n/2$  tels axes.

Notons que cette fois  $G$  ne contient plus que  $n/2$  symétries passant par un sommet et un côté. L'ordre de  $G$  est donc encore  $2n$ . La formule de Burnside donne

$$\text{Card}(\Omega) = \frac{1}{2n} \left( \sum_{d|n} \varphi(d) q^{n/d} + \frac{n}{2} q^{(n+2)/2} + \frac{n}{2} q^{n/2} \right). \quad (18.500)$$

### 18.10.15 Classification

THOooRORQooTDWFdv

**Théorème 18.188** ([605]).

Toute isométrie du plan  $(\mathbb{R}^2, d)$  est une composition d'au plus 3 réflexions.

*Démonstration.* Encore une fois nous décomposons la preuve en fonction du nombre de points fixes.

- (1) **Si  $f$  n'a pas de point fixe** Soit  $x \in \mathbb{R}^2$ . Nous considérons le segment  $[x, f(x)]$  et nous nommons  $l$  sa médiatrice. Par construction,  $f(x) = \sigma_l(x)$ . Nous posons  $g = \sigma_l \circ f$ , et nous avons

$$g(x) = x. \quad (18.501)$$

Donc nous avons  $f = \sigma_l \circ g$  avec  $x \in \text{Fix}(g)$ .

- (2) **Si  $f$  a un unique point fixe** Soit  $x$  cet unique point fixe. Soit  $y \neq x$  et  $l$  la médiatrice de  $[y, f(y)]$ . En posant  $g = \sigma_l \circ f$  nous avons

$$g(y) = y \quad (18.502)$$

et  $g(x) = x$  parce que

$$d(x, f(y)) = d(f(x), f(y)) = d(x, y), \quad (18.503)$$

ce qui donne que  $x$  est à égale distance de  $y$  et de  $f(y)$ , c'est-à-dire que  $x \in l$  et par conséquent  $g(x) = (\sigma_l \circ f)(x) = \sigma_l(x) = x$ .

Donc  $g$  fixe  $x$  et  $y$ . Par conséquent,  $g$  fixe toute la droite  $(xy)$ .

- (3) **Si  $f$  fixe une droite** Soit  $l$  une droite fixée par  $f$ . Alors  $f$  est soit  $\sigma_l$  soit l'identité par le lemme 18.114.

- (4) **Conclusion**

Nous avons montré que si  $\text{Fix}(f)$  a pour dimension  $m$ , alors il existe une droite pour laquelle  $f = \sigma_l \circ g$  avec  $\dim(\text{Fix}(g)) > m$ . Donc il faut au maximum trois pas pour avoir  $\dim(\text{Fix}(g)) = 2$  c'est-à-dire pour avoir  $g = \text{Id}$ .

□

DEFooJEOYooNwYtuQ

**Définition 18.189.**

Une **réflexion glissée** est une transformation du plan de la forme  $\tau_v \circ \sigma_\ell$  où le vecteur  $v$  est parallèle à la droite  $\ell$ .

THOooVRNOooAgaVRN

**Théorème 18.190** ([605]).

Les isométries du plan  $(\mathbb{R}^2, d)$  sont exactement

- (1) l'identité (composée de 0 réflexions),
- (2) les réflexions,
- (3) les translations (composées de 2 réflexions d'axes parallèles),
- (4) les rotations (composées de 2 réflexions d'axes non parallèles),

(5) *les réflexions glissées (composées de 3 réflexions)*

*Démonstration.* Nous savons déjà que  $f \in \text{Isom}(\mathbb{R}^2)$  est une composée de 0, 1, 2 ou 3 réflexions.

- (1) **Zéro réflexion** Alors c'est l'identité. Ce n'est pas très profond.
- (2) **Une réflexion** Alors  $f$  est une réflexion. Toujours pas très profond.
- (3) **Deux réflexions** Soit  $f = \sigma_{\ell_1} \circ \sigma_{\ell_2}$ . Maintenant ça s'approfondit un bon coup.

Nous supposons d'abord que  $\ell_1 \parallel \ell_2$ . Dans ce cas nous allons prouver que  $f = \tau_{2v}$  où  $v$  est le vecteur perpendiculaire à  $\ell_1$  tel que  $\ell_1 + v = \ell_2$ . Nous allons utiliser le lemme 18.113 pour montrer que  $\sigma_{\ell_1} \circ \sigma_{\ell_2} = \tau_{2v}$ . Nous avons

$$\ell_1 = \ell_0 + w \quad (18.504a)$$

$$\ell_2 = \ell_0 + w + v \quad (18.504b)$$

où  $w$  est un vecteur perpendiculaire à  $\ell_1$  et  $\ell_0$  est la droite passant par l'origine et parallèle à  $\ell_1$  et  $\ell_2$ . Avec cela,

$$(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x) = \sigma_{\ell_1}(\sigma_{\ell_0}(x) + 2w) \quad (18.505a)$$

$$= \sigma_{\ell_0}(\sigma_{\ell_0}(x) + 2w) + 2(v + w) \quad (18.505b)$$

$$= x + \underbrace{\sigma_{\ell_0}(2w)}_{-2w} + 2v + 2w \quad (18.505c)$$

$$= x + 2v. \quad (18.505d)$$

Donc si  $f$  est composée de deux réflexions d'axes parallèles, alors  $f$  est une translation.

Toujours dans le cas où  $f$  est composée de deux réflexions, nous supposons que  $f = \sigma_{\ell_2} \circ \sigma_{\ell_1}$  avec  $\ell_1$  et  $\ell_2$  non parallèles. Nous notons  $O$  le point d'intersection, et nous allons voir que  $f = R_O(2\alpha)$ , où  $\alpha$  est l'angle de  $\ell_1$  à  $\ell_2$  donné par le lemme 18.161.

Soit  $x \in \ell_1$ . Alors

$$f(x) = \sigma_{\ell_2}(x), \quad (18.506)$$

et le lemme 18.164 nous donne un moyen de calculer  $\sigma_{\ell_2}(x)$  parce que  $\ell_1$  est une droite passant par  $x$  et coupant  $\ell_2$  au point  $O$ . Le lemme dit que  $\sigma_{\ell_2}(x) = R_O(2\alpha)$ . Remarque : c'est bien  $2\alpha$  et non  $-2\alpha$  parce qu'il s'agit de l'angle de  $\ell_2$  à  $\ell_1$  ; il y a inversion des numéros entre ici et l'énoncé du lemme.

Nous avons donc bien  $f(x) = R_O(2\alpha)x$  pour  $x \in \ell_1$ .

Si  $y \in \ell_2$  alors

$$f(y) = \sigma_{\ell_2}(R_O(-2\alpha)y) \quad (18.507)$$

Nous posons  $z = \sigma_{\ell_1}(y) = R_O(-2\alpha)y$ . Soit la droite  $\ell_3$  passant par  $O$  et  $z$ . Puisque  $R_O(2\alpha)z = y \in \ell_2$ , l'angle de  $\ell_3$  à  $\ell_2$  est  $2\alpha$ . Par conséquent

$$\sigma_{\ell_2}(z) = R_O(-2 \times (-2\alpha))z = R_O(4\alpha)z = R_O(4\alpha)R_O(-2\alpha)y = R_O(2\alpha)y. \quad (18.508)$$

Donc les transformations  $f$  et  $R_O(2\alpha)$  coïncident pour tous les points des droites  $\ell_1$  et  $\ell_2$ , qui ne sont pas parallèles. Cela prouve que  $f = R_O(2\alpha)$ .

- (4) **Trois réflexions** Nous écrivons  $f = \sigma_{\ell_3} \circ \sigma_{\ell_2} \circ \sigma_{\ell_1}$ . Nous allons transformer cela progressivement en une symétrie glissée en passant par plusieurs étapes :

$$(4a) \quad f = \sigma_{\ell} \circ \tau_v, \quad \text{ITEMooHVYCooPhFMiv}$$

$$(4b) \quad f = \tau_v \circ \sigma_{\ell}, \quad \text{ITEMooUKGLooFlCejt}$$

$$(4c) \quad f = \tau_v \circ \sigma_{\ell} \text{ avec } v \parallel \ell. \quad \text{ITEMooWUCWooZSjofe}$$

À chacune de ces étapes,  $v$  et  $\ell$  vont changer. La dernière est une réflexion glissée.

Nous commençons par supposer  $\ell_2 \parallel \ell_3$ . Dans ce cas,  $\sigma_{\ell_3} \circ \sigma_{\ell_2}$  est une translation, comme nous l'avons déjà vu. Alors  $f = \tau_v \circ \sigma_{\ell_1}$  et nous sommes déjà dans le cas (4b).

Nous supposons que  $\ell_2$  n'est pas parallèle à  $\ell_3$ . Dans ce cas, si  $O = \ell_2 \cap \ell_3$  nous avons

$$\sigma_{\ell_3} \circ \sigma_{\ell_2} = R_O(2\alpha) \quad (18.509)$$

où  $\alpha$  est l'angle de  $\ell_2$  à  $\ell_3$ . En réalité tant que l'angle de  $\ell'_2$  à  $\ell'_3$  est  $\alpha$  nous avons

$$\sigma_{\ell'_3} \circ \sigma_{\ell'_2} = \sigma_{\ell_3} \circ \sigma_{\ell_2} = R_O(2\alpha). \quad (18.510)$$

Nous choisissons  $\ell'_2$  parallèle à  $\ell_1$ , de telle sorte que  $\sigma_{\ell'_2} \circ \sigma_{\ell_1}$  soit une translation. Alors nous avons

$$f = \sigma_{\ell_3} \circ \sigma_{\ell_2} \circ \sigma_{\ell_1} = \sigma_{\ell'_3} \circ \sigma_{\ell'_2} \circ \sigma_{\ell_1} = \sigma_{\ell_3} \circ \tau_v. \quad (18.511)$$

où  $v$  est le vecteur de la translation en question.

Nous avons donc prouvé que toute composition de trois réflexions peut être écrite soit sous la forme (4a) soit sous la forme (4b).

Nous prouvons à présent que toute transformation de la forme (4a) peut être écrite sous la forme (4b). Plus précisément nous allons prouver que si  $\ell$  est une droite,  $v$  un vecteur et  $\ell_0$  la droite parallèle à  $\ell$  passant par l'origine, alors

$$\sigma_\ell \circ \tau_v = \tau_{\sigma_{\ell_0}(v)} \circ \sigma_l \quad (18.512)$$

D'abord nous savons que  $\sigma_\ell(x) = \sigma_{\ell_0}(x) + 2w$  où  $w$  est le vecteur tel que  $\ell = \ell_0 + w$ . Ensuite c'est un simple calcul utilisant le fait que  $\sigma_{\ell_0}$  est linéaire :

$$(\sigma_\ell \circ \tau_v)(x) = \sigma_l(x + v) = \sigma_{\ell_0}(x) + \sigma_{\ell_0}(v) + 2w, \quad (18.513)$$

et

$$(\tau_{\sigma_{\ell_0}(v)} \circ \sigma_\ell)(x) = \sigma_{\ell_0}(v) + \sigma_\ell(x) = \sigma_{\ell_0}(v) + \sigma_{\ell_0}(x) + 2w. \quad (18.514)$$

L'égalité est prouvée.

Nous montrons maintenant que toute transformation de la forme (4b) peut être mise sous la forme (4c). Soit donc  $f = \tau_v \circ \sigma_\ell$  où  $v$  et  $\ell$  ne sont pas spécialement parallèles.

Pour cela nous décomposons  $v = v_1 + v_2$  avec  $v_1 \perp \ell$  et  $v_2 \parallel \ell$  et nous posons  $\ell' = \ell + \frac{1}{2}v_1$ . Nous montrons que

- $\tau_v \circ \sigma_\ell = \tau_{v_2} \circ \sigma_{\ell'}$
- $v_2 \parallel \ell'$ .

Pour le deuxième point,  $v_2 \parallel \ell$  et bien entendu  $\ell' \parallel \ell$ . Donc  $v_2 \parallel \ell'$ .

Soit  $\ell_0$  la droite parallèle à  $\ell$  et  $\ell'$  et passant par l'origine. Soit aussi le vecteur  $w$  tel que  $\ell = \ell_0 + w$ . Alors nous avons

$$\begin{cases} \sigma_\ell = \sigma_{\ell_0} + 2w & (18.515a) \\ \sigma_{\ell'} = \sigma_{\ell_0} + 2w + v_1 & (18.515b) \end{cases}$$

Nous avons

$$(\tau_v \circ \sigma_\ell)(x) = v + \sigma_{\ell_0}(x) + 2w \quad (18.516)$$

et

$$(\tau_{v_2} \circ \sigma_{\ell'})(x) = v_2 + \sigma_{\ell_0}(x) + 2w + v_1 \quad (18.517a)$$

$$= \sigma_{\ell_0}(x) + v + 2w \quad (18.517b)$$

où dans la dernière ligne, nous avons regroupé  $v_1 + v_2 = v$ . Et voilà.

□

**18.10.16 Classification des isométries de  $\mathbb{R}$** **Définition 18.191.**

Soit  $x \in \mathbb{R}$  ; nous notons  $\sigma_x$  la **réflexion** par rapport à  $x$ , c'est-à-dire

$$\sigma_x(y) = 2x - y. \quad (18.518)$$

**Théorème 18.192 ([605]).**

Toute isométrie de  $\mathbb{R}$  est composée d'au plus 2 réflexions. Plus précisément toute isométrie de  $\mathbb{R}$  est dans une des trois catégories suivantes :

- l'identité (0 réflexion),
- les réflexions,
- les translations (2 réflexions)

*Démonstration.* Nous divisons la preuve en fonction du nombre de points fixés par l'isométrie  $f \in \text{Isom}(\mathbb{R})$ .

- (1)  **$f$  fixe deux points distincts** Alors elle fixe l'espace affine engendré par ces deux points par la proposition 18.102. Donc  $f$  fixe tout  $\mathbb{R}$ , et est l'identité.
- (2)  **$f$  fixe un unique point** Soit  $x$  l'unique point fixé par  $f$  et considérons  $y \neq x$ . Puisque  $x = f(x)$  et que  $f$  est une isométrie,

$$d(x, f(y)) = d(f(x), f(y)) = d(x, y). \quad (18.519)$$

Donc  $f(y)$  est à égale distance de  $x$  que de  $y$ . Autrement dit,  $f(y)$  est, soit  $y$ , soit  $\sigma_x(y)$ . Mais comme  $x$  est unique point fixe,  $f(y) = \sigma_x(y)$ . Ce raisonnement étant valable pour tout  $y \neq x$ , nous avons  $f = \sigma_x$ .

- (3)  **$f$  n'a pas de point fixe** Soient  $x \in \mathbb{R}$  et  $y = \frac{x+f(x)}{2}$ . Nous posons  $g = \sigma_y \circ f$ . Alors  $x$  est un point fixe de  $g$  parce que

$$g(x) = \sigma_y(f(x)) = 2y - f(x) = x. \quad (18.520)$$

Donc, soit  $g$  est l'identité, soit  $g$  est une réflexion (par les points précédents). La possibilité  $g = \text{Id}$  est exclue parce que cela donnerait  $f = \sigma_y$  alors que  $f$  n'a pas de point fixe. Donc  $g$  est une réflexion ; et comme  $x$  est un point fixe de  $g$  nous avons  $g = \sigma_x$ . Au final

$$f = \sigma_y \circ \sigma_x. \quad (18.521)$$

Montrons que cela implique que  $f$  est une translation :

$$\sigma_y \sigma_x(z) = \sigma_y(2x - z) = 2y - 2x + z = z + 2(y - x). \quad (18.522)$$

Donc  $\sigma_y \circ \sigma_x$  est la translation de vecteur  $2(y - x)$ .

□

**18.10.17 Isométries du tétraèdre régulier**

PROPOoHTBIooTJQNZt

**Proposition 18.193.**

Soient un tétraèdre régulier<sup>52</sup>  $T$  ainsi qu'une application  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  qui vérifie

- (1)  $f$  est affine<sup>53</sup>
- (2)  $f$  est bijective (sur  $\mathbb{R}^3$ )
- (3)  $f(T) \subset T$ .

Alors

<sup>52</sup>. Définition 12.180

<sup>53</sup>. Définition 8.156.



- (1) Il existe une bijection  $\alpha: \{1, 2, 3, 4, 5, 6\} \rightarrow \{1, 2, 3, 4, 5, 6\}$  telle que  $f(d_i) = d_{\alpha(i)}$ .  
 (2) Pour la même bijection  $\alpha: \{1, 2, 3, 4, 5, 6\} \rightarrow \{1, 2, 3, 4, 5, 6\}$  nous avons  $f(a_i) = a_{\alpha(i)}$ .  
 (3) Il existe une bijection  $\beta: \{1, 2, 3, 4\} \rightarrow \{1, 2, 3, 4\}$  telle que  $f(s_i) = s_{\beta(i)}$ .

Autrement dit,  $f$  agit comme une permutation<sup>54</sup> sur l'ensemble des sommets de  $T$ .

*Démonstration.* Puisque  $d_i$  est une droite et que  $f$  est affine, l'ensemble  $f(d_i)$  est une droite de  $\mathbb{R}^3$  (proposition 12.145).

- (1) **Pour (1)** Puisque  $f(T) \subset T$ , la partie  $f(a_i)$  contient une infinité de points alignés dans  $T$  (parce que  $a_i$  contient une infinité de points et  $f$  est une bijection de  $\mathbb{R}^3$ ).

Comme le tétraèdre n'a que 6 arêtes, il y a forcément une des arêtes qui contient au moins deux points de  $f(a_i)$ . Soit  $a_{\beta(i)}$  une arête qui contient deux points de  $f(a_i)$ . En particulier  $d_{\alpha(i)}$  contient deux points de  $f(a_i)$ , et donc de  $f(d_i)$ . Donc  $f(d_i) = d_{\alpha(i)}$ .

- (2) **Pour (2)** L'ensemble  $f(a_i)$  est contenu dans  $d_{\alpha(i)} \cap T = a_{\alpha(i)}$  (lemme 12.182). Nommons  $s_1$  et  $s_2$  les deux sommets de  $T$  sur  $a_i$ . Soit  $k \in \{1, 2\}$ . Comme les sommets sont sur plusieurs arêtes, il existe  $j \neq i$  tel que  $s_k = a_i \cap a_j$ . Nous avons alors

$$f(s_k) \in f(a_i) \cap f(a_j) \subset a_{\alpha(i)} \cap a_{\alpha(j)}. \quad (18.523)$$

Mais nous avons déjà vu que les seules intersections des segments sont les sommets (lemme 12.184). Donc  $f(s_k)$  est un sommet de  $T$ .

Nous avons prouvé que  $f(a_i)$  contient deux sommets sur  $a_{\alpha(i)}$  donc il contient tout  $a_{\alpha(i)}$ . Cela prouve que  $f(a_i) = a_{\alpha(i)}$ .

- (3) **Pour (3)** Nous avons déjà mentionné, juste en-dessous de (18.523) que l'image d'un sommet doit être un sommet. Nous avons donc une application  $\beta: \{1, 2, 3, 4\} \rightarrow \{1, 2, 3, 4\}$  telle que  $f(s_i) = s_{\beta(i)}$ . Puisque  $f$  est injective,  $\beta$  est injective. Par principe des tiroirs, cette application doit également être surjective.

□

PROPooVNLKooOjQzCj

**Proposition 18.194** (Isométries affines du tétraèdre régulier).

Soient  $T$  un tétraèdre régulier et  $\text{Isom}(T)$  son groupe d'isométries affines<sup>55</sup>. Alors

$$\text{Isom}(T) \simeq S_4 \quad (18.524)$$

où  $S_4$  est le groupe des permutations de quatre objets.

*Démonstration.* Si  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  est une bijection affine (ce qui est le cas d'une isométrie affine), alors la proposition 18.193(3) donne une bijection  $\beta: \{1, 2, 3, 4\} \rightarrow \{1, 2, 3, 4\}$  des sommets de  $T$ . Si nous numérotions les sommets  $x_1, \dots, x_4$ , nous obtenons un morphisme de groupe  $\varphi: \text{Isom}(T) \rightarrow S_4$  qui envoie  $g$  sur la permutation qui envoie 1 sur le numéro du sommet  $g(x_1)$ , 2 sur le numéro du sommet  $g(x_2)$ , etc.

- (1) **Le morphisme  $\varphi$  est injectif** Supposons  $\varphi(g_1) = \varphi(g_2)$ . Alors  $g_1^{-1} \circ g_2$  est une isométrie de  $(\mathbb{R}^3, d)$  qui fixe les quatre sommets. Une application affine  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$  fixant 4 points est l'identité par la proposition 8.158. Donc  $g_1^{-1} \circ g_2 = \text{Id}$ , ce qui prouve que  $g_1 = g_2$ . Vous noterez que nous utilisons l'unicité de l'inverse dans un groupe.

- (2)  **$\varphi$  est surjectif**

Nous savons que  $S_4$  est engendré par les transpositions (proposition 1.283). Or les transpositions sont dans l'image de  $\varphi$ . En effet, notons les sommets de notre tétraèdre par  $A, B, C$  et  $D$  et considérons la transposition  $A \leftrightarrow B$ . Elle est l'image par  $\varphi$  de la réflexion selon le plan  $\sigma$ , médiateur du segment  $[A, B]$ . Pour nous assurer de cela, nous devons nous assurer que  $C$  et  $D$  appartiennent à  $\sigma$ . C'est le contenu du lemme 12.185.

- (3) **Conclusion** L'application  $\varphi$  est un morphisme bijectif, c'est-à-dire un isomorphisme.

□

54. Une permutation est une bijection, définition 1.261.

55. Une isométrie est une isométrie, et une application est affine dans la définition 8.156.

18.10.18 Représentation de  $S_4$  via les isométries du tétraèdre

SUBSECooVEASooDUbsBh

## 18.195.

Lorsque le tétraèdre a son barycentre en l'origine de  $\mathbb{R}^3$ , l'isomorphisme  $\varphi: \text{Isom}(T) \rightarrow S_4$  donne une représentation de dimension 3 de  $S_4$ . Nous avons calculé les caractères de  $S_4$  en la section 16.5 sans avoir besoin de savoir que l'une des représentations de dimension 3 est celle que nous venons de trouver via le groupe des isométries du tétraèdre. Nous allons cependant également y calculer les caractères de la représentation  $\varphi$ , pour le sport.

Une des représentations trouvées (la représentation  $\rho_s$ ) peut être vue comme le groupe  $\text{Isom}(T)$  des isométries affine du tétraèdre grâce à la proposition 18.194 qui donne un isomorphisme de groupe  $S_4 \simeq \text{Iso}(T)$  lorsque  $T$  est un tétraèdre régulier de  $\mathbb{R}^3$ .

Si le barycentre de  $T$  est situé à l'origine de  $\mathbb{R}^3$ , alors les éléments de  $\text{Isom}(T)$  sont des applications linéaires parce que

- les affinités laissent invariantes les barycentres (proposition 11.43),
- les affinités qui laissent l'origine invariante sont linéaires (corolaire 11.60).

Nous allons à présent calculer la trace de cette représentation, en utilisant le fait que nous la connaissions explicitement. Nous savons que les caractères sont constants sur les classes de conjugaison; nous allons donc écrire une matrice par classe de conjugaison (qui sont données dans l'exemple 1.279).

Pour tout cela nous allons considérer un tétraèdre dont l'isobarycentre est en  $(0, 0, 0)$  et une base de  $\mathbb{R}^3$  formée de trois sommets  $e_1, e_2$  et  $e_3$ . Puisque l'isobarycentre des quatre sommets est en  $(0, 0, 0)$ , le quatrième sommet est forcément le point de coordonnées  $e_4(-1, -1, -1)$ , de telle sorte que  $e_1 + e_2 + e_3 + e_4 = 0$ .

**Les transpositions** Quelle isométrie de  $\mathbb{R}^3$  permute deux sommets du tétraèdre sans bouger les autres? Pour permuter les sommets  $e_1$  et  $e_2$  en laissant  $e_3$  et  $e_4$ , c'est la symétrie par rapport au plan médiateur de  $[e_1, e_2]$ . Ce plan passe par les sommets  $e_3$  et  $e_4$ , parce que le tétraèdre étant régulier, les points  $e_3$  et  $e_4$  sont équidistants de  $e_1$  et  $e_2$ . Le lemme 12.185 dit qu'alors, ces points font partie du plan médiateur.

Dans notre base, la matrice de la transposition précédemment nommée (12) est

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (18.525)$$

dont la trace est 1. Donc  $\chi_s(12) = 1$ .

**Les bitranspositions** La bitransposition (12)(34) est le produit des transpositions selon les plans médiateurs de  $[e_1, e_2]$  et  $[e_3, e_4]$ . Ces deux plans sont perpendiculaires, et l'intersection est la droite qui passe par les milieux. Cette droite est perpendiculaire aux deux segments en même temps. La matrice est :

$$\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} \quad (18.526)$$

parce que  $e_1 \mapsto e_2, e_2 \mapsto e_1$  et  $e_3 \mapsto e_4$ . Pour rappel, la matrice est formée des images des vecteurs de base. Cela donne

$$\chi_s((12)(34)) = -1. \quad (18.527)$$

**Les 3-cycles** La symétrie qui permute cycliquement les points  $e_1, e_2$  et  $e_3$  est la rotation d'angle <sup>56</sup>  $2\pi/3$  dans le plan formé par les extrémités de ces trois vecteurs. Heureusement, la trace est invariante par changement de base; donc nous pouvons calculer la trace d'une rotation d'angle

56. Angle d'une rotation, définition 18.126.

$2\pi/3$  dans n'importe quelle base. Par exemple :

$$\chi_s((12)(34)) = \text{Tr} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2\pi/3) & \sin(2\pi/3) \\ 0 & -\sin(2\pi/3) & \cos(2\pi/3) \end{pmatrix} = 1 + 2 \cos(2\pi/3) = 0. \quad (18.528)$$

Notons que, sans cette interprétation géométrique, nous y arrivons aussi facilement : dans notre base le 3-cycle est  $e_1 \mapsto e_2 \mapsto e_3 \mapsto e_1$ , donc la matrice est :

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad (18.529)$$

dont la trace est manifestement nulle :  $\chi_s((123)) = 0$ .

**Le 4-cycle** Il réalise  $e_1 \mapsto e_2 \mapsto e_3 \mapsto e_4 \mapsto e_1$ , dont la matrice est

$$\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \quad (18.530)$$

et la trace est  $\chi_s((1, 2, 3, 4)) = -1$ .

Nous avons retrouvé les caractères de la représentation  $\rho_s$ , et nous pouvons vérifier qu'elle est irréductible.

## 18.11 Transformations de Lorentz

Nous considérons dans cette section un nombre réel  $c > 0$  ainsi que l'espace  $\mathbb{R}^2$  muni du produit pseudo-scalaire<sup>57</sup> donné par la matrice

$$\eta = \begin{pmatrix} c^2 & 0 \\ 0 & -1 \end{pmatrix}. \quad (18.531)$$

Et pour faire plus vrai, nous notons  $(x_0, x_1)$  les coordonnées sur  $\mathbb{R}^2$ . Ainsi

$$x \cdot y = c^2 x_0 y_0 - x_1 y_1. \quad (18.532)$$

Nous insistons sur le fait que cela n'est pas un produit scalaire.

**Lemme 18.196** ([1]).

Soit  $c > 0$ . L'application

$$\begin{aligned} \varphi: ]-c, c[ &\rightarrow \mathbb{R} \\ v &\mapsto \frac{-v/c}{\sqrt{1 - \frac{v^2}{c^2}}} \end{aligned} \quad (18.533)$$

est une bijection.

*Démonstration.* Nous commençons par mentionner le fait que  $\varphi$  est continue du fait que le dénominateur ne s'annule pas. Une petite étude de fonction montre que

$$\lim_{v \rightarrow -c} \varphi(v) = +\infty, \quad (18.534)$$

et

$$\lim_{v \rightarrow c} \varphi(v) = -\infty, \quad (18.535)$$

---

<sup>57</sup>. Définition 8.165.

et

$$\varphi'(v) = -\frac{1}{c\sqrt{1-\frac{v^2}{c^2}}} - \frac{v^2/c^3}{\left(1-\frac{v^2}{c^2}\right)^{3/2}} < 0. \quad (18.536)$$

Tout cela fait que  $\varphi$  est bijective (entre autres par le théorème des valeurs intermédiaires 9.95 et le théorème dérivée et croissance 12.209).  $\square$

LEMooUZFKooSIjery

### Lemme 18.197.

La forme bilinéaire

$$\begin{aligned} b: \mathbb{R}^2 \times \mathbb{R}^2 &\rightarrow \mathbb{R} \\ x, y &\mapsto x \cdot y \end{aligned} \quad (18.537)$$

est non dégénérée<sup>58</sup>.

*Démonstration.* Soit  $(x_0, x_1) \in \mathbb{R}^2$  tel que

$$b((x_0, x_1), (y_0, y_1)) = 0 \quad (18.538)$$

pour tout  $(y_0, y_1) \in \mathbb{R}^2$ . Nous avons

$$c^2 x_0 y_0 - x_1 y_1 = 0. \quad (18.539)$$

En écrivant cela avec  $(y_0, y_1) = (1, 0)$  puis  $(0, 1)$  nous obtenons immédiatement que  $(x_0, x_1) = (0, 0)$ .  $\square$

THOooYHDWooWxVovH

### Théorème 18.198.

Soit une bijection<sup>59</sup>  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  telle que

$$f(x) \cdot f(y) = x \cdot y \quad (18.540)$$

pour tout  $x, y \in \mathbb{R}^2$ . Alors :

(1)  $f$  est linéaire.

(2) Il existe un unique choix de  $(\xi, \sigma_1, \sigma_2) \in \mathbb{R} \times \{\pm 1\} \times \{\pm 1\}$  tel que la matrice de  $f$  ait la forme

$$f = \begin{pmatrix} \sigma_1 \cosh(\xi) & \frac{\sigma_1 \sigma_2}{c} \sinh(\xi) \\ c \sinh(\xi) & \sigma_2 \cosh(\xi) \end{pmatrix}. \quad (18.541)$$

(3) Il existe un unique  $v \in ]-c, c[$  tel que la matrice de  $f$  ait la forme

$$f = \begin{pmatrix} \frac{\sigma_1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\sigma_1 \sigma_2}{c^2} \frac{v}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{-v}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{\sigma_2}{\sqrt{1-\frac{v^2}{c^2}}} \end{pmatrix}. \quad (18.542)$$

*Démonstration.* Puisque notre produit pseudo-scalaire est non dégénéré (lemme 18.197), le fait que  $f$  soit linéaire est la proposition 8.154. Nous posons

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad (18.543)$$

et, conformément à la proposition 8.159, nous imposons  $A^t \eta A = \eta$ . Après un petit produit matriciel nous obtenons :

$$\begin{pmatrix} c^2 \alpha^2 - \gamma^2 & c^2 \alpha \beta - \gamma \delta \\ c^2 \alpha \beta - \gamma \delta & c^2 \beta^2 - \delta^2 \end{pmatrix} = \begin{pmatrix} c^2 & 0 \\ 0 & -1 \end{pmatrix}. \quad (18.544)$$

58. Définition 8.126.

59. À mon avis, il y a moyen d'affaiblir cette hypothèse. Écrivez-moi si vous avez une idée.

Voilà quatre équations à résoudre pour les quatre inconnues  $\alpha, \beta, \gamma, \delta$ . Déjà les équations des termes anti-diagonaux sont les mêmes. Nous recopions le reste :

$$\begin{cases} c^2 \alpha^2 - \gamma^2 = c^2 & (18.545a) \\ c^2 \alpha \beta - \gamma \delta = 0 & (18.545b) \\ c^2 \beta^2 - \delta^2 = 1 & (18.545c) \end{cases}$$

C'est le moment d'utiliser la proposition 15.142. La relation (18.545a) donne

$$\alpha^2 - \left(\frac{\gamma}{c}\right)^2 = 1, \quad (18.546)$$

ce qui implique l'existence d'un unique <sup>60</sup>  $\xi_1 \in \mathbb{R}$  et  $\sigma_1 \in \{\pm 1\}$  tels que

$$\gamma = c \sinh(\xi_1) \quad (18.547a)$$

$$\alpha = \sigma_1 \cosh(\xi_1). \quad (18.547b)$$

La relation (18.545c) implique quant à elle l'existence de  $\xi_2 \in \mathbb{R}$  et  $\sigma_2 \in \{\pm 1\}$  tels que

$$\delta = \sigma_2 \cosh(\xi_2) \quad (18.548a)$$

$$\beta = \frac{1}{c} \sinh(\xi_2). \quad (18.548b)$$

Nous substituons maintenant toutes les valeurs (18.547) et (18.548) dans (18.545b). Cela donne

$$\sigma_1 \cosh(\xi_1) \sinh(\xi_2) = \sigma_2 \sinh(\xi_1) \cosh(\xi_2). \quad (18.549)$$

Nous mettons cette relation au carré et nous substituons  $\cosh(\xi_1)^2 = 1 + \sinh^2(\xi_1)$ . Ce que nous trouvons est

$$\sinh(\xi_1)^2 = \sinh(\xi_2)^2, \quad (18.550)$$

qui implique que  $\xi_1 = \pm \xi_2$ . Nous posons donc  $\xi_2 = \sigma_3 \xi_1$  pour un certain  $\sigma_3 \in \{\pm 1\}$ . Cela nous permet d'alléger la notation et d'écrire  $\xi$  au lieu de  $\xi_1$ .

Nous remettons la valeur  $\xi = \xi_1 = \sigma_3 \xi_2$  dans l'équation (18.549) en tenant compte du fait que  $\sinh$  est impaire et  $\cosh$  est paire :

$$\sigma_1 \sigma_3 \cosh(\xi) \sinh(\xi) = \sigma_2 \sinh(\xi) \cosh(\xi). \quad (18.551)$$

Et cela nous enseigne que  $\sigma_3 = \sigma_1 \sigma_2$ .

Jusqu'à présent nous avons prouvé qu'il existe un unique  $\xi \in \mathbb{R}$  et  $\sigma_1, \sigma_2 \in \{\pm 1\}$  tels que

$$A = \begin{pmatrix} \sigma_1 \cosh(\xi) & \frac{\sigma_1 \sigma_2}{c} \sinh(\xi) \\ c \sinh(\xi) & \sigma_2 \cosh(\xi) \end{pmatrix}. \quad (18.552)$$

Nous utilisons à présent la bijection du lemme 18.196. Il existe un unique  $v \in ]-c, c[$  tel que  $\sinh(\xi) = \varphi(v)$ . En utilisant  $\cosh(\xi)^2 = 1 + \varphi(v)^2$ , nous trouvons

$$\cosh(\xi)^2 = \frac{1}{1 - \frac{v^2}{c^2}}. \quad (18.553)$$

Mais comme le cosinus hyperbolique est toujours strictement positif, nous pouvons prendre la racine carrée des deux côtés :

$$\cosh(\xi) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (18.554)$$

En substituant dans (18.552), nous trouvons le résultat annoncé.  $\square$

---

60. Par 15.142.

## 18.11.1 Sous-groupe fini d'isométries du plan

**Théorème 18.199** ([617]).Soit un groupe fini  $G$  d'isométries de  $(\mathbb{R}^2, d)$  contenant  $n$  éléments.

- (1) Il existe un point  $C \in \mathbb{R}^2$  fixé par tous les éléments de  $G$ .
- (2) Si  $G$  ne contient pas de réflexion, alors il est cyclique<sup>61</sup> et engendré par la rotation d'angle  $2\pi/n$  autour de  $C$ .
- (3) Si  $G$  contient au moins une réflexion, et si  $C$  est un point fixe de  $G$ , alors
  - (3a) toutes les réflexions ont un axe qui passe par  $C$ ,
  - (3b)  $n$  est pair,
  - (3c) Si  $\sigma$  est une réflexion dans  $G$ , alors nous avons  $G = \text{gr}(\sigma, R_C(4\pi/n))$  où  $R_C(\theta)$  est la rotation d'angle  $\theta$  autour de  $C$ ,
  - (3d)  $G$  est isomorphe au groupe diédral  $D_{n/2}$ .

*Démonstration.* Soit un groupe fini  $G$  constitué d'isométries de  $(\mathbb{R}^2, d)$ . Nous prouvons le théorème point par point.

- (1) **Pour (1)** C'est la proposition 18.99.
- (2) **Questions de réflexions** Le théorème 18.188 nous dit que les éléments de  $G$  sont des compositions d'au maximum 3 réflexions.
- (3) **Exclure trois réflexions** Il n'est pas possible que  $G$  contienne un élément composé de trois réflexions. En effet, les composées de trois réflexions, par le théorème 18.190 sont des réflexions glissées<sup>62</sup>, c'est-à-dire des transformations de la forme  $g = \tau_v \circ \sigma_\ell$  où  $v$  est un vecteur parallèle à la droite  $\ell$ . Si  $x \in \ell$ , alors

$$g(x) = \tau_v(x) = x + v, \quad (18.555)$$

de telle sorte que  $g^k(x) = x + kv$ , qui signifie que tous les  $g^k$  sont différents. Le groupe  $G$  ne peut pas être fini si il contient une réflexion glissée.

- (4)  **$G^+$  et  $G^-$**  Pour la même raison que celle qui exclut les réflexions glissées,  $G$  ne peut pas contenir de translation. Le théorème 18.190 nous donne la liste des possibilités. Après exclusion des translations et des réflexions glissées, il reste :
  - l'identité,
  - les rotations,
  - les réflexions.

Nous notons  $G^+$  la partie de  $G$  contenant l'identité et les rotations et  $G^-$  celle contenant les réflexions. Notons que  $G^+$  n'est pas vide parce qu'il contient au moins l'identité, tandis que  $G^-$  peut être vide, mais n'est certainement pas un groupe.

- (5) **Même nombre d'éléments** Nous prouvons à présent que si  $G^-$  est non vide, alors il a le même nombre d'éléments que  $G^+$ . Un élément de  $G^-$  est une réflexion. Soit  $\sigma \in G^-$ . Nous prouvons que

$$\begin{aligned} \varphi: G^+ &\rightarrow G^- \\ f &\mapsto \sigma \circ f \end{aligned} \quad (18.556)$$

est une bijection.

- (5a) **Surjective** Soit  $s \in G^-$ . Posons  $f = \sigma^{-1} \circ s$ . Puisque  $\sigma^{-1}$  et  $s$  sont des réflexions,  $f$  est une rotation. Donc  $f \in G^+$  et  $\varphi(f) = s$ .
- (5b) **Injective** La condition  $\varphi(f) = \varphi(g)$  dit que  $\sigma \circ f = \sigma \circ g$ . En composant par  $\sigma^{-1}$  nous obtenons  $f = g$ .

61. Définition 1.428.

62. Définition 18.189.

- (6)  $G^+ = \text{gr}(R_C(2\pi/p))$  Nous nommons  $p$  le nombre d'éléments de  $G^+$ . Si  $G^-$  est vide,  $p = n$ , et sinon  $p = n/2$ . Dans les deux cas,  $G^+$  est un groupe de rotations à  $p$  éléments.

Le groupe  $G^+$  contient seulement des rotations ; or le centre d'une rotation est l'unique point fixe. Donc tous les éléments de  $G^+$  sont des rotations autour de  $C$ .

Le corolaire 2.15 au théorème de Lagrange nous indique que tous les éléments de  $G^+$  vérifient  $g^p = \text{Id}$ . Seules les rotations d'angle  $2k\pi/p$  autour de  $C$  satisfont la condition  $g^p = \text{Id}$ . Or il n'y a que  $p$  telles rotations. Donc elles sont toutes dans  $G^+$ . Nous en déduisons que

$$G^+ = \text{gr}(R_C(2\pi/p)). \quad \text{EQooUWTVooEMokVH} \quad (18.557)$$

- (7) **Pour (2)** Dans le cas où  $G$  ne contient pas de réflexion,  $G^-$  est vide et  $G$  contient  $n$  éléments. La relation (18.557) devient

$$G = G^+ = \text{gr}(R_C(2\pi/n)). \quad (18.558)$$

- (8) **Pour (3)** Nous supposons maintenant que  $G$  contienne au moins une réflexion. De la sorte  $G^- \neq \emptyset$ .

- (8a) **Pour (3a)** Les seuls points fixes d'une réflexion sont ceux de l'axe. Donc  $C$  doit être sur tous les axes des réflexions contenues dans  $G^-$ .

Notons au passage que deux réflexions d'axes qui se coupent forment une rotation. Donc  $G^-$  ne forme pas un groupe, mais même pas en rêve, quoi.

- (8b) **Pour (3b)** Vu que l'union  $G = G^+ \cup G^-$  est disjointe et que  $G^+$  et  $G^-$  ont le même nombre d'éléments par la bijection 18.556, si  $G^-$  est non vide,  $G$  possède un nombre pair d'éléments.

- (8c) **Pour (3c)** Si  $\sigma \in G$  est une réflexion, nous savons que  $n$  est pair, que  $G^+$  possède  $p = n/2$  éléments et que

$$G^+ = \{R_C(2k\pi/p)\} = \{R_C(4k\pi/n)\}_{k=1,\dots,n/2}. \quad (18.559)$$

L'élément  $\sigma \in G^-$  étant fixé, la bijection (18.556) nous indique que tous les éléments de  $G^-$  sont de la forme  $\sigma \circ f$  avec  $f \in G^+$ . Donc

$$G^- \subset \text{gr}(\sigma, R_C(4\pi/n)). \quad (18.560)$$

Nous avons aussi

$$G^+ \subset \text{gr}(\sigma, R_C(4\pi/n)). \quad (18.561)$$

Et comme  $\sigma$  et  $R_C(4\pi/n)$  sont dans  $G$  nous avons  $\text{gr}(\sigma, R_C(4\pi/n)) \subset G$ . Tout cela pour dire que

$$G = \text{gr}(\sigma, R_C(4\pi/n)). \quad (18.562)$$

- (8d)  $R\sigma = \sigma R^{-1}$  Nous restons dans le cas où  $G^-$  n'est pas vide. Nous considérons  $R$ , la rotation d'angle  $\theta$  autour de  $C$ . Si  $R_0$  est la rotation d'angle  $\theta$  autour de  $(0,0)$ , nous avons

$$R = \tau_C \circ R_0 \circ \tau_C^{-1}, \quad (18.563)$$

et si  $\sigma_0$  est la symétrie d'axe parallèle à l'axe de  $\sigma$ , mais passant par  $(0,0)$  nous avons :

$$\sigma = \tau_C \circ \sigma_0 \circ \tau_C^{-1}. \quad (18.564)$$

Si  $v$  est le vecteur directeur de la réflexion  $\sigma_0$ , nous considérons enfin  $\alpha$ , la rotation qui réalise  $\alpha(v) = (1,0)$ . Nous avons alors

$$\sigma_0 = \alpha^{-1} \circ s \circ \alpha \quad (18.565)$$

où  $s$  est la symétrie autour de l'axe horizontal. En n'ayant pas peur d'identifier  $\mathbb{R}^2$  à  $\mathbb{C}$ , l'application  $s$  est la conjugaison complexe. Avec tout ça nous avons

$$R\sigma = \tau_C R_0 \tau_C^{-1} \tau_C \sigma_0 \tau_C^{-1} = \tau_C R_0 \alpha^{-1} s \alpha \tau_C^{-1} = \tau_C \alpha^{-1} R_0 s \alpha \tau_C^{-1} \quad (18.566)$$

où nous avons utilisé le fait que les rotations autour de  $(0, 0)$  forment un groupe abélien pour commuter  $\alpha^{-1}$  avec  $R_0$ . Nous utilisons à présent le lemme 18.177 pour commuter  $R$  avec  $s$  :

$$R\sigma = \tau_C \alpha^{-1} s R_0^{-1} \alpha \tau_C^{-1} \quad (18.567a)$$

$$= \tau_C \underbrace{\alpha^{-1} s \alpha}_{\sigma_0} R_0^{-1} \tau_C^{-1} \quad (18.567b)$$

$$= \tau_C \sigma_0 \tau_C^{-1} \tau_C R_0^{-1} \tau_C^{-1} \quad (18.567c)$$

$$= \sigma R^{-1}. \quad (18.567d)$$

Nous avons utilisé le fait que  $\tau_C R_0^{-1} \tau_C^{-1} = R^{-1}$  comme on peut s'en convaincre en calculant le produit.

(8e) **Table de multiplication**

Nous considérons une réflexion  $\sigma \in G$ . Les éléments de  $G^+$  sont des rotations autour de  $C$  et ceux de  $G^-$  de la forme  $\sigma R$  où  $R$  est une rotation autour de  $C$ . Pour connaître la table de multiplication de  $G$ , nous devons écrire

$$(\sigma^{\epsilon_1} R^k)(\sigma^{\epsilon_2} R^l) = \sigma^{\epsilon} R^m \quad (18.568)$$

où  $\epsilon_1, \epsilon_2 \in \{0, 1\}$ ,  $R$  est la rotation d'angle  $4\pi/n$  autour de  $C$  et  $\epsilon$  et  $m$  sont des constantes à exprimer en fonction de  $\epsilon_1, \epsilon_2, k$  et  $l$ .

Tous les éléments de  $G$  pouvant être écrits soit sous la forme  $R^m$ , soit sous la forme  $\sigma R^m$ , nous avons les possibilités suivantes :

$$(8e.i) \quad R^m R^l = R^{m+l}$$

$$(8e.ii) \quad (R^m)(\sigma R^l) = \sigma R^{-m} R^l = \sigma R^{l-m}$$

$$(8e.iii) \quad (\sigma R^m) R^l = \sigma R^{m+l}$$

$$(8e.iv) \quad (\sigma R^m)(\sigma R^l) = \sigma \sigma R^{l-m} = R^{l-m}.$$

(8f) **Pour (3d)** Récoltons quelque acquis.

- Nous venons de prouver que  $R\sigma = \sigma R^{-1}$ .
- Tout élément de  $G$  peut s'écrire soit sous la forme  $R^m$ , soit sous la forme  $\sigma R^m$ , selon que l'élément est dans  $G^+$  ou  $G^-$ .
- Tout élément du groupe diédral  $D_n$  s'écrit soit sous la forme  $r^m$ , soit sous la forme  $sr^m$  (proposition 18.174(2)).

L'application  $\varphi: G \rightarrow D_n$  suivante est donc une bijection :

$$\left\{ \begin{array}{l} \varphi(R^m) = r^m \\ \varphi(\sigma R^m) = sr^m. \end{array} \right. \quad (18.569a)$$

$$\quad \quad \quad (18.569b)$$

Il nous reste à prouver que c'est un morphisme. Cela se fait en utilisant la table de multiplication du groupe diédral donnée dans la proposition 18.178 et celle du groupe  $G$  que nous venons d'écrire.

□

**Définition 18.200** (Groupe de symétrie d'une partie de  $\mathbb{R}^n$  [605]).

Si  $Y$  est une partie de  $\mathbb{R}^n$ , nous définissons le **groupe des symétries** de  $Y$  par

$$\text{Sym}(Y) = \{f \in \text{Isom}(\mathbb{R}^n) \text{ tel que } f(Y) = Y\}. \quad (18.570)$$

Nous définissons aussi le groupe des symétries propres de  $Y$  par

$$\text{Sym}^+(Y) = \{f \in \text{Isom}^+(\mathbb{R}^n) \text{ tel que } f(Y) = Y\} \quad (18.571)$$

où  $\text{Isom}^+(\mathbb{R}^n)$  est le groupe des isométries positives de  $\mathbb{R}^n$ , définition 18.93.



**Théorème 18.201** ([605]).

Soit  $Y \subset \mathbb{R}^2$ .

- (1) Si le groupe  $\text{Sym}^+(Y)$  est fini d'ordre  $n$ , alors c'est un groupe cyclique d'ordre  $n$ .
- (2) Si  $\text{Sym}^+(Y)$  est fini, alors  $\text{Sym}(Y)$  est soit cyclique<sup>63</sup> d'ordre  $n$ , soit isomorphe au groupe diédral<sup>64</sup> d'ordre  $2n$ .

*Démonstration.* Nous savons déjà par la proposition 18.100 que  $\text{Sym}^+(Y)$  est isomorphe à un sous-groupe  $H^+$  d'ordre  $n$  de  $\text{SO}(2)$ . Vérifions que ce groupe est cyclique. Si  $n = 1$ , c'est évident. Si  $n \geq 2$  alors nous savons que  $H^+$  est constitué de rotations d'angles dans  $[0, 2\pi[$  et vu que c'est un ensemble fini, il possède une rotation d'angle minimal (à part zéro). Notons  $\alpha_0$  cet angle.

Nous montrons que  $H^+$  est engendré par la rotation d'angle  $\alpha_0$ . Soit une rotation d'angle  $\alpha$ . Étant donné que  $\alpha_0 < \alpha$  nous pouvons effectuer la division euclidienne<sup>65</sup> de  $\alpha$  par  $\alpha_0$  et obtenir

$$\alpha = k\alpha_0 + \beta \quad (18.572)$$

avec  $\beta < \alpha_0$ . Mézalors  $R(\beta) = R(\alpha)R(\alpha_0)^{-k}$  est également un élément du groupe. Cela contredit la minimalité dès que  $\beta \neq 0$ . Avoir  $\beta = 0$  revient à dire que  $\alpha$  est un multiple de  $\alpha_0$ , ce qui signifie que le groupe  $H^+$  est cyclique engendré par  $\alpha_0$ .

Notons au passage que nous avons automatiquement  $\alpha_0 = \frac{2\pi}{n}$  parce qu'il faut  $R(\alpha_0)^n = \text{Id}$ . Nous avons prouvé que  $\text{Sym}^+(Y)$  est cyclique d'ordre  $n$ .

Nous étudions maintenant le groupe  $\text{Sym}(Y)$ . Par la proposition 18.100 nous avons un morphisme injectif

$$\phi: \text{Sym}(Y) \rightarrow \text{O}(2), \quad (18.573)$$

et en posant  $H = \phi(\text{Sym}(Y))$  nous avons un isomorphisme de groupes  $\phi: \text{Sym}(Y) \rightarrow H$ . Nous savons aussi que ce  $\phi$  se restreint en

$$\phi: \text{Sym}^+(Y) \rightarrow H^+ \subset \text{SO}(2) \quad (18.574)$$

où  $H^+ = \phi(\text{Sym}^+(Y)) = H \cap \text{SO}(2)$ . Le groupe  $H^+$  est cyclique et est engendré par la rotation  $R(2\pi/n)$ .

Supposons un instant que  $H \subset \text{SO}(2)$ . Alors nous avons  $H = H^+$  et  $\phi$  est un isomorphisme entre  $\text{Sym}(Y)$  et le groupe cyclique engendré par  $R(2\pi/n)$ .

Nous supposons à présent que  $H$  n'est pas un sous-ensemble de  $\text{SO}(2)$ . Quelles sont les isométries de  $\mathbb{R}^2$  qui ne sont pas de déterminant 1 ? Il faut regarder dans le théorème 18.190 quelles sont les isométries contenant un nombre impair de réflexions. Ce sont les réflexions et les réflexions glissées. Or il ne peut pas y avoir de réflexion glissée dans un groupe fini, parce que si  $f$  est une réflexion glissée, tous les  $f^k$  sont différents.

Nous en déduisons que si  $H$  n'est pas inclus dans  $\text{SO}(2)$ , il contient une réflexion que nous nommons  $\sigma$ . Nous allons en déduire que  $H \simeq H^+ \times_{\mathbf{Ad}} C_2$  où  $C_2 = \{\text{Id}, \sigma\}$ . Si  $h \in H$  nous pouvons écrire

$$h = (h\sigma^\epsilon)\sigma^\epsilon \quad (18.575)$$

pour n'importe quelle valeur de  $\epsilon$ , et en particulier pour  $\epsilon = \pm 1$ .

Si  $h \in \text{SO}(2)$  alors nous écrivons  $h = h\sigma^0$  et si  $h \notin \text{SO}(2)$  nous écrivons  $h = (h\sigma)\sigma$ . Vu que  $h\sigma \in \text{SO}(2)$ , cette dernière écriture est encore de la forme  $\text{SO}(2) \times C_2$ . Quoi qu'il en soit, tout élément de  $H$  s'écrit comme un produit

$$H = H^+ C_2. \quad (18.576)$$

Cette décomposition est unique parce que si  $h_1 c_1 = h_2 c_2$  alors  $h_2^{-1} h_1 = c_2 c_1^{-1}$ , et comme  $h_2^{-1} h_1 \in H^+$  nous avons  $c_2 c_1^{-1} \in H^+$  et donc  $c_1 = c_2$ . Partant, nous avons aussi  $h_1 = h_2$ . Pour avoir

63. Définition 1.428.

64. Définition 18.169.

65. Théorème 1.343.

le produit semi-direct, il faut encore montrer que  $\mathbf{Ad}(C_2)H^+ \subset H^+$ . Le seul cas à vérifier est  $\mathbf{Ad}(\sigma)H^+ \subset H^+$ . Comme les éléments de  $H^+$  sont caractérisés par le fait d'avoir un déterminant positif, nous avons

$$\mathbf{Ad}(\sigma)R(\alpha) = \sigma R(\alpha)\sigma^{-1} \in H^+. \quad (18.577)$$

□

### Remarque 18.202.

Tout ceci est cohérent avec le théorème de Burnside 8.311 parce que le sous-groupe fini de  $\mathrm{SO}(n)$  engendré par la rotation  $R(2\pi/n)$  est un groupe d'exposant fini, à savoir que si  $h$  est dans ce groupe,  $h^n = \mathrm{Id}$ .

## 18.11.2 Relations trigonométriques dans un triangle rectangle

Nous donnons maintenant quelques relations trigonométriques classiques dans un triangle rectangle. Le théorème de Pythagore est déjà le théorème 10.33; nous nous concentrons ici sur les angles.

PROPOCDZVooKOQzct

### Proposition 18.203.

Soient  $A, B, S \in \mathbb{R}^2$  des points distincts et non alignés formant un triangle rectangle en  $A$  :

$$(A - S) \cdot (B - A) = 0. \quad (18.578)$$

En posant  $\theta = \widehat{ASB}$  nous avons

$$\cos(\theta) = \frac{\|A - S\|}{\|B - S\|} \quad (18.579)$$

et

$$\sin(\theta) = \pm \frac{\|B - A\|}{\|B - S\|}. \quad (18.580)$$

EQooEKZFEooFeNImX

*Démonstration.* Nous posons  $C = A - S$  et  $D = B - S$ . Comme  $C \neq 0$ , il existe une rotation  $R_\alpha$  telle que

$$\begin{cases} (R_\alpha C)_x > 0 \\ (R_\alpha C)_y = 0. \end{cases} \quad (18.581a)$$

$$(18.581b)$$

Nous posons  $X = R_\alpha C$  et  $Y = R_\alpha D$ .

Le triangle formé de  $O$ ,  $X$  et  $Y$  est « posé » sur l'axe des abscisses et est rectangle en  $X$ , c'est-à-dire

$$X \cdot (Y - X) = 0. \quad (18.582)$$

De ce fait, le point  $Y$  satisfait à  $Y_x = X_x$ . Et enfin, grâce aux propositions 18.162 et 18.163 nous avons  $\widehat{ASB} = \widehat{XOY}$ .

Nous écrivons les relations qui définissent l'angle  $\widehat{XOY}$ . Pour cela nous posons  $X' = X/\|X\|$  et  $Y' = Y/\|Y\|$  et nous avons <sup>SUBEQooVHNDooPOfbjC</sup>

$$\begin{cases} \cos(\theta) = X'_x Y'_x \\ \sin(\theta) = X'_y Y'_y. \end{cases} \quad (18.583a)$$

$$(18.583b)$$

Vu que  $X = (X_x, 0)$ , nous avons  $X'_x = 1$ . De plus

$$\|Y\| = \|R_\alpha(D)\| = \|D\| = \|B - S\|. \quad (18.584)$$

En substituant les valeurs dans (18.583),

$$\cos(\theta) = Y'_x = \frac{Y_x}{\|Y\|} = \frac{X_x}{\|B - S\|} = \frac{\|C\|}{\|B - S\|} = \frac{\|A - S\|}{\|B - S\|}. \quad (18.585)$$

Voilà déjà une chose de prouvée.

Pour la seconde, nous avons  $\sin(\theta) = Y'_y$ . Selon le signe de  $Y_y$  nous avons  $Y_y = \pm \|Y - X\|$  et donc

$$\sin(\theta) = Y'_y = \frac{Y_y}{\|Y\|} = \frac{\pm \|Y - X\|}{\|Y\|} = \pm \frac{\|D - C\|}{\|Y\|} = \pm \frac{\|B - A\|}{\|B - S\|}. \quad (18.586)$$

□

Le signe sur la formule du sinus revient au fait que la définition de l'angle  $\widehat{AOB}$  est de considérer la rotation qui fait aller  $A$  vers  $B$ . Donc suivant la position relative de  $A$ ,  $O$  et  $B$ , il se peut que l'angle mesuré soit l'angle *extérieur* au triangle.

La proposition suivante est parfois prise comme définition de l'angle.

**Proposition 18.204.**

Soient trois points non alignés  $A, S, B \in \mathbb{R}^2$ . Nous avons

$$\cos(\widehat{ASB}) = \frac{(A - S) \cdot (B - S)}{\|A - S\| \|B - S\|}. \quad (18.587)$$

*Démonstration.* Nous posons  $C = A - S$ ,  $D = B - S$ ,  $X = C/\|C\|$  et  $Y = D/\|D\|$ . Avec cela, la définition 18.156 donne l'équation

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} X_x \\ X_y \end{pmatrix} = \begin{pmatrix} Y_x \\ Y_y \end{pmatrix} \quad (18.588)$$

que nous écrivons comme le système

$$\begin{cases} X_x \cos(\theta) - X_y \sin(\theta) = Y_x \\ X_y \cos(\theta) + X_x \sin(\theta) = Y_y. \end{cases} \quad (18.589a)$$

$$(18.589b)$$

Nous considérons maintenant cela comme un système pour  $(\cos(\theta), \sin(\theta))$  :

$$\begin{pmatrix} X_x & -X_y \\ X_y & X_x \end{pmatrix} \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} Y_x \\ Y_y \end{pmatrix}. \quad (18.590)$$

Le déterminant de la dernière matrice est  $X_x + X_y^2 = \|X\|^2 = 1$  parce que  $X$  est unitaire. Cette matrice est donc inversible et son inverse est vite calculée. Nous avons

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} X_x & X_y \\ -X_y & X_x \end{pmatrix} \begin{pmatrix} Y_x \\ Y_y \end{pmatrix}. \quad (18.591)$$

Cela donne ce que nous voulions :

$$\cos(\theta) = X \cdot Y = \frac{C \cdot D}{\|C\| \|D\|} = \frac{(A - S) \cdot (B - S)}{\|A - S\| \|B - S\|}. \quad (18.592)$$

□

**Remarque 18.205.**

Prendre la formule (18.587)

$$\cos(\widehat{ASB}) = \frac{(A - S) \cdot (B - S)}{\|A - S\| \|B - S\|}. \quad (18.593)$$

comme définition de l'angle  $\widehat{ASB}$  est cependant trompeur parce que ça ne permet de définir les angles que sur une partie de  $[0, 2\pi[$  sur laquelle le cosinus est injectif. Pour réellement définir tous les angles, il faut alors un peu bricoler.

Sans vouloir être méchant, je crois que ceux qui procèdent de cette manière sont ceux qui donnent un cours sur le produit scalaire sans avoir l'intention de lier la définition d'une rotation comme composée de réflexions aux matrices de  $SO(2)$  et aux fonctions trigonométriques.

### 18.11.3 Pavages du plan

DEFooHPKFooSIDhCM

#### Définition 18.206.

Une application affine  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  est un **déplacement** lorsqu'elle est une isométrie de  $(\mathbb{R}^n, d)$  qui préserve l'orientation<sup>66</sup>.

DEFooJPHKooRgCBJs

#### Définition 18.207 ([486]).

Un **pavage** de  $\mathbb{R}^2$  est une paire  $(G, K)$  où  $G$  est un groupe de déplacements<sup>67</sup> de  $\mathbb{R}^2$  et  $K$  un compact de  $\mathbb{R}^2$  d'intérieur non vide telle que

$$(1) \quad G \cdot K = \mathbb{R}^2,$$

ITEMooOIJZooZMKLUm

$$(2) \quad \text{Si } g_1, g_2 \in G \text{ satisfont } g_1 \cdot \text{Int}(K) \cap g_2 \cdot \text{Int}(K) \neq \emptyset, \text{ alors } g_1 \cdot K = g_2 \cdot K.$$

Nous disons qu'un groupe  $G$  de déplacements de  $\mathbb{R}^2$  est un **groupe de pavage** de  $\mathbb{R}^2$  si il existe un compact  $K$  tel que la paire  $(G, K)$  soit un pavage.

En termes de notations,

$$G \cdot K = \bigcup_{g \in G} g(K) = \bigcup_{g \in G} \bigcup_{k \in K} g(k). \quad (18.594)$$

LEMooWZSWooZYkiCn

#### Lemme 18.208 ([1]).

Soient une bijection affine  $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  ainsi qu'une droite  $d$  et un point  $A$ . Nous notons  $S$  la partie de  $\mathbb{R}^2$  située du côté de  $d$  contenant  $A$ .

Alors  $\varphi(S)$  est la partie de  $\mathbb{R}^2$  située du côté de  $\varphi(d)$  contenant  $\varphi(A)$ .

*Démonstration.* La droite  $d$  est donnée par une application affine  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  et la définition

$$d = \{x \in \mathbb{R}^2 \text{ tel que } f(x) = 0\}. \quad (18.595)$$

Nous supposons que  $f(A) > 0$ ; sinon, nous pouvons utiliser  $-f$  au lieu de  $f$ . Donc

$$S = \{x \in \mathbb{R}^2 \text{ tel que } f(x) > 0\}. \quad (18.596)$$

La partie  $\varphi(S)$  est alors donnée par

$$\varphi(S) = \{\varphi(x) \text{ tel que } f(x) > 0\}. \quad (18.597)$$

Comme  $\varphi$  est une bijection, cela s'écrit aussi bien

$$\varphi(S) = \{y \in \mathbb{R}^2 \text{ tel que } f(\varphi^{-1}(y)) > 0\}. \quad (18.598)$$

De même

$$\varphi(d) = \{s \in \mathbb{R}^2 \text{ tel que } f(\varphi^{-1}(s)) = 0\}. \quad (18.599)$$

Donc les deux côtés de la droite  $\varphi(d)$  sont donnés par le signe de  $f \circ \varphi^{-1}$ . Nous avons

$$(f \circ \varphi^{-1})(\varphi(A)) = f(A) > 0. \quad (18.600)$$

Donc  $\varphi(A) \in \varphi(S)$ . □

LEMooZOXVooTJiLTF

#### Lemme 18.209.

Le groupe

$$G = \text{gr}(\tau_{e_1}, \tau_{e_2}) \quad (18.601)$$

est un groupe de pavage du plan.

*Démonstration.* Il suffit de prendre le carré  $K = [0, 1] \times [0, 1]$ . En appliquant les translations, nous recouvrons tout le plan, sans intersection des intérieurs des carrés. Notons toutefois qu'il y a un recouvrement des bords. □

<sup>66</sup>. Définition 8.34.

<sup>67</sup>. Définition 18.206.

LEMooTMRGooChBzZg

**Lemme 18.210** ([1, 618]).*Le groupe*

$$G = \text{gr}(\tau_{e_1}, \tau_{e_2}, R_0(\pi)) \quad (18.602)$$

*est un groupe de pavage du plan.*

*Démonstration.* Le compact à considérer est  $K = [0, \frac{1}{2}] \times [0, 1]$ . Le compact  $K$  et son image par  $R_0(\pi)$  sont représentés sur la figure 18.8.

En agissant sur  $K$  avec les translations verticales et horizontales, nous recouvrons des bandes verticales de largeur  $1/2$ . En agissant de même sur  $R_0(\pi)(K)$ , nous recouvrons les autres bandes verticales.

Donc  $G \cdot K$  recouvre bien  $\mathbb{R}^2$ . Il serait cependant un peu présomptueux de croire en avoir fini. Il faut vérifier la condition (2) de la définition 18.207 d'un pavage.

Supposons que  $g_1 \cdot \text{Int}(K) \cap g_2 \cdot \text{Int}(K) \neq \emptyset$ . Cela signifie qu'il existe  $k_1, k_2 \in \text{Int}(K)$  tels que  $g_1(k_1) = g_2(k_2)$ , ou encore que

$$(g_2^{-1}g_1)k_1 = k_2 \in \text{Int}(K). \quad (18.603)$$

EQoolTJEoolUkUKuu

Quelle est la forme d'un élément général de  $G$ ? Le lemme 1.425 nous indique qu'un élément général de  $G$  est un produit fini de  $\tau_{e_1}$ ,  $\tau_{e_2}$  et  $R_0(\pi)$ . Mais nous savons que si  $\alpha$  est linéaire,

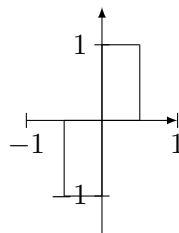
$$\alpha \circ \tau_u = \tau_{\alpha(u)} \circ \alpha. \quad (18.604)$$

Dans notre cas, dans un produit général, nous pouvons déplacer tous les facteurs  $R_0(\pi)$  à droite en changeant des  $\tau_{e_i}$  en  $\tau_{R_0(\pi)e_i} = \tau_{-e_i}$ . Les translations par contre commutent sans faire d'histoires. Donc un élément général de  $G$  est de la forme

$$g = \tau_{e_1}^k \tau_{e_2}^l R_0(\pi)^m \quad (18.605)$$

Nous pouvons évidemment restreindre  $m$  à  $\{0, 1\}$ . Supposons  $k \in \text{Int}(K)$  et  $g(k) \in \text{Int}(K)$ . Nous avons  $0 < k_x < 0.5$ . Si  $m = 1$ , alors  $g(k)_x \in ]-1/2, 0[$  et aucun  $\tau_{e_1}^k g(k)_x$  ne pourra plus être entre 0 et  $1/2$ . Donc  $m = 0$ . À partir de là, pour avoir  $g(k) \in \text{Int}(K)$  nous devons avoir également  $k = l = 0$ . Donc  $g = \text{Id}$ .

Deux éléments  $g_1$  et  $g_2$  vérifiant la condition (18.603) doivent donc vérifier  $g_2^{-1}g_1 = \text{Id}$ , et donc  $g_1 = g_2$ . Par conséquent  $g_1 \cdot K = g_2 \cdot K$ .  $\square$

FIGURE 18.8: Le compact  $K$  et son image par  $R_0(\pi)$  pour le lemme 18.210

LabelFigATJSooefYkmCbP

LEMooJPNDoohDCLnY

**Lemme 18.211** ([1, 618]).*Le groupe*

$$\text{gr}(\tau_{e_1}, \tau_{e_2}, R_0(\pi/2)) \quad (18.606)$$

*est un groupe de pavage.*

*Démonstration.* Le pavé à considérer est

$$K = [-\frac{1}{2}, 0] \times [0, \frac{1}{2}]. \quad (18.607)$$

En lui appliquant trois fois la rotation  $R_0(\pi/2)$ , nous reconstituons le carré  $[-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$ . Ensuite, avec les translations, nous pavons tout le plan.

Pour la seconde condition, nous procédons comme dans la démonstration du lemme 18.210. D'abord  $R_0(\pi/2)e_1 = e_2$  et  $R_0(\pi/2)e_2 = -e_1$ . Donc dans un produit général de  $\tau_{e_1}$ ,  $\tau_{e_2}$  et  $R_0(\pi/2)$  (et de leurs inverses), toutes les rotations peuvent être mises à droite ; nous avons donc un élément général de  $G$  sous la forme

$$g = \tau_a \circ R_0(\pi/2)^k \quad (18.608)$$

avec  $a \in \mathbb{Z}e_1 + \mathbb{Z}e_2$  et  $k \in \{0, 1, 2, 3\}$ .

Vu que les translations se font par nombres entiers tandis que les différences de coordonnées entre les  $R_0(\pi/2)^k K$  sont demi-entiers, si  $k \in \text{Int}(K)$ , alors aucun  $\tau_a$  ne permet d'avoir  $(\tau_a \circ R_0(\pi/2))^k \in \text{Int}(K)$ .

Bref, si  $(g_2^{-1}g_1)k \in K$ , alors encore une fois  $g_2^{-1}g_1 = \text{Id}$ . □

Et c'est maintenant que les choses compliquées commencent.

LEMoMWWEEooEbZXtb

### Lemme 18.212.

*Le groupe*

$$\text{gr} \left( \tau_{e_1}, \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(\pi/3) \right) \quad (18.609)$$

*est un groupe de pavage.*

*Démonstration.* Ici le compact  $K$  est le triangle de sommets  $A = (0, 0)$ ,  $B = (\frac{1}{2}, \frac{\sqrt{3}}{6})$  et  $C = (1, 0)$ . Plus précisément il s'agit de l'intersection des trois parties suivantes :

- le côté de la droite  $(AB)$  où est  $C$ ,
- le côté de la droite  $(AC)$  où est  $B$ ,
- le côté de la droite  $(BC)$  où est  $A$ .

Il est bon d'écrire ces trois conditions sous forme d'inéquations.

- La droite  $AB$  est donnée par l'équation  $f_{AB}(x, y) = 0$  pour  $f_{AB}(x, y) = x - \sqrt{3}y$ . Puisque  $f_{AB}(C) = 1$ , la première inéquation pour  $K$  est

$$x - \sqrt{3}y \geq 0. \quad (18.610)$$

- Pour la droite  $(BC)$  nous avons  $f_{BC}(x, y) = -x - \sqrt{3}y + 1$  et  $f_{BC}(A) = 1$ . Donc la seconde inéquation pour  $K$  est

$$-x - \sqrt{3}y + 1 \geq 0 \quad (18.611)$$

- La droite  $AC$  est donnée par l'application  $f_{AC}(x, y) = y$ . Vu que  $f_{AC}(B) = \sqrt{3}/6$ , nous avons la troisième inéquation pour  $K$  :

$$y \geq 0. \quad (18.612)$$

En résumé, la définition de  $K$  est le système SUBEQSooECKFooOdneOA

$$\begin{cases} x - \sqrt{3}y \geq 0 & (18.613a) \\ -x - \sqrt{3}y + 1 \geq 0 & (18.613b) \\ y \geq 0. & (18.613c) \end{cases}$$

La figure 18.9(a) nous montre ce triangle et l'action des puissances de  $R_0(\pi/3)$  sur lui. Pour votre gouverne, la matrice de cette rotation est

$$R_0(\pi/3) = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}. \quad (18.614) \quad \text{EQooMRBXooTabLit}$$

La figure 18.9(c) vous montre une partie de ce pavage dans toute sa splendeur.

Puisque les translations sont  $u_1 = e_1$  et  $u_2 = (\frac{1}{2}, \frac{\sqrt{3}}{2})$ , il est suffisant de montrer que notre pavage pave réellement le parallélogramme construit sur ces deux vecteurs. Nous vous avons très obligeamment dessiné ce parallélogramme pavé sur la figure 18.9(b).

Pour montrer que les triangles dessinés pavent effectivement le parallélogramme, nous allons procéder à une exhaustion de cas. Soit  $(x, y)$  dans le parallélogramme.

Nous commençons par couper le parallélogramme en deux parties suivant la diagonale allant de  $u_1$  à  $u_2$ . Dans la vie mon p'ti gars, il y a deux types de points : ceux qui vérifient  $x + \frac{1}{\sqrt{3}}y - 1 \leq 0$  et les autres.

(1) Si  $x + y/\sqrt{3} - 1 \leq 0$  Ceci est le côté de  $(0, 0)$ . Nous subdivisons suivant la petite barre verticale (suivez le dessin), c'est-à-dire suivant les deux cas :  $x \geq \frac{1}{2}$  et  $x \leq \frac{1}{2}$ .

(1a) Si  $x \leq \frac{1}{2}$  Enfin nous coupons avec la droite diagonale partant de  $(0, 0)$ , c'est-à-dire selon que  $x - \sqrt{3}y \leq 0$  ou  $x - \sqrt{3}y \geq 0$ .

(1a.i) Si  $x - \sqrt{3}y \leq 0$  Les points dont nous parlons sont les  $(x, y) \in \mathbb{R}^2$  vérifiant |<sup>SUBEQSooNYWDooYMNvad</sup>

$$\begin{cases} x + \frac{1}{\sqrt{3}}y - 1 \leq 0 \end{cases} \quad (18.615a)$$

$$\begin{cases} x \leq \frac{1}{2} \end{cases} \quad (18.615b)$$

$$\begin{cases} x - \sqrt{3}y \leq 0. \end{cases} \quad (18.615c)$$

En suivant le dessin, vous remarquerez que l'élément de  $G$  à considérer est

$$g = \tau_{e_1} \circ \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})} \circ R_0(\pi/3)^4. \quad (18.616)$$

Nous avons |<sup>EQSooOJBfooCTaTtu</sup>

$$g(A) = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) \quad (18.617a)$$

$$g(B) = \left(\frac{1}{2}, \frac{\sqrt{3}}{6}\right) \quad (18.617b)$$

$$g(C) = (0, 0). \quad (18.617c)$$

Ce que les équations (18.615) décrivent est l'intersection des trois parties suivantes<sup>68</sup> :

- le côté de la droite  $(g(A)g(B))$  contenant  $g(C)$ ,
- le côté de la droite  $(g(A)g(C))$  contenant  $g(B)$ ,
- le côté de la droite  $(g(B)g(C))$  contenant  $g(A)$ ,

Le lemme 18.208 nous permet d'exprimer ces trois parties en termes de  $K$  :

- l'image par  $g$  du côté de la droite  $(AB)$  contient  $C$ ,
- l'image par  $g$  du côté de la droite  $(AC)$  contient  $B$ ,
- l'image par  $g$  du côté de la droite  $(BC)$  contient  $A$ ,

Comme  $g$  est une bijection, l'intersection des images par  $g$  est l'image par  $g$  de l'intersection. Bref, les équations (18.617) décrivent l'image par  $g$  de  $K$ .

Cette partie est donc pavée par  $(G, K)$ .

(1a.ii) Si  $x - \sqrt{3}y \geq 0$  hop, même calculs avec des petites variations. Nous laissons la lectrice s'en occuper.

(1b) Si  $x \geq \frac{1}{2}$  hop.

(2) Si  $x + y/\sqrt{3} - 1 \geq 0$  hop.

Les cas listés se traitent sûrement de la même façon. Nous tenons pour prouvé que  $(G, K)$  est bien surjectif sur  $\mathbb{R}^2$ .

Nous devons encore montrer la condition (2) de la définition 18.207. Commençons par déterminer la forme générale d'un élément de  $G$  sous la forme  $\tau_a \circ r_0$  où  $r_0$  est une application linéaire. Le

68. Personnellement, je n'ai pas vérifié, mais ça m'étonnerait que ce soit faux. Vérifiez et écrivez-moi.

lemme 1.425 nous indique qu'un élément général de  $G = \text{gr}(\tau_{e_1}, \tau_{(\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(\pi/3))$  est un produit arbitraire (mais fini) de  $\tau_{e_1}$ ,  $\tau_{(\frac{1}{2}, \frac{\sqrt{3}}{2})}$  et de  $R_0(\pi/3)$  et de leurs inverses.

Comme toujours nous avons  $\alpha \circ \tau_v = \tau_{\alpha(v)} \circ \alpha$ . Donc nous pouvons passer toutes les rotations  $R_0(\pi/3)$  et  $R_0(\pi/3)^{-1}$  à droite du produit, quitte à produire des translations des formes suivantes :

$$R_0(\pi/3)^k e_1 \quad (18.618a)$$

$$R_0(\pi/3)^k \left( \frac{1}{2}, \frac{\sqrt{3}}{2} \right) \quad (18.618b)$$

$$R_0(\pi/3)^{-k} e_1 \quad (18.618c)$$

$$R_0(\pi/3)^{-k} \left( \frac{1}{2}, \frac{\sqrt{3}}{2} \right). \quad (18.618d)$$

En utilisant la matrice (18.614) nous trouvons assez vite que <sup>SUBEQooEMVlooNaaMqk</sup>

$$R_0(\pi/3) e_1 = \left( \frac{1/2}{\sqrt{3}/2} \right) \quad (18.619a)$$

$$R_0(\pi/3)^2 e_1 = \left( \frac{-1/2}{\sqrt{3}/2} \right) \quad (18.619b)$$

$$R_0(\pi/3)^3 e_1 = -e_1. \quad (18.619c)$$

Les applications suivantes de  $R_0(\pi/3)$  ne donnent rien de nouveau, si ce n'est le signe. Les puissances de  $R_0(\pi/3)$  appliquées à  $(\frac{1}{2}, \frac{\sqrt{3}}{2})$  sont déjà parmi celles listées en (18.619). Quant aux inverses,  $R_0(\pi/3)^{-1} = R_0(\pi/3)^5$ ; donc rien de nouveau non plus.

Un élément général de  $G$  est donc dans

$$\tau_v \circ R_0(\pi/3)^k \quad (18.620)$$

avec  $v$  dans

$$\mathbb{Z}e_1 + \mathbb{Z} \left( \frac{1/2}{\sqrt{3}/2} \right) + \mathbb{Z} \left( \frac{-1/2}{\sqrt{3}/2} \right). \quad (18.621)$$

Remarquons que

$$\left( \frac{-1/2}{\sqrt{3}/2} \right) = -e_1 + \left( \frac{1/2}{\sqrt{3}/2} \right). \quad (18.622)$$

Donc un élément général de  $G$  est

$$\tau_{e_1}^k \circ \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}^l \circ R_0(\pi/3)^m. \quad (18.623)$$

Nous pouvons maintenant prouver notre point. Pour cela nous allons successivement considérer les 5 rotations de  $K$  présentées dans la sous-figure 18.9(a). Pour chacune nous allons montrer qu'aucune translation ne permet d'obtenir une intersection avec  $K$ .

Nous allons en faire un seul en détail.

- (1) **Pour**  $L = R_0(\pi/3)K$  hop. les détails sont laissés à notre aimable lecteur qui ne pourra pas ne pas compléter tous les cas.
- (2) **Pour**  $L = R_0(\pi/3)^2 K$  Nous allons prouver que si  $(x, y) \in \text{Int}(L)$ , alors

$$\tau_{e_1}^k \circ \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}^l (x, y) \quad (18.624)$$

ne peut pas être dans  $\text{Int}(K)$ . Pour la simplicité des notations nous notons  $r = R_0(\pi/3)^2$ ; nous avons

$$r = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}. \quad (18.625)$$



Nous considérons  $g = \tau_{e_1}^k \circ \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}^l \circ r$ . Un calcul nous donne l'image de  $(x, y)$  par  $g$  :

$$g(x, y) = \tau_{e_1}^k \circ \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}^l \left( \begin{pmatrix} -\frac{1}{2}x - \frac{\sqrt{3}}{2}y \\ -\frac{\sqrt{3}}{2}x - \frac{1}{2}y \end{pmatrix} \right) = \begin{pmatrix} -\frac{1}{2}x - \frac{\sqrt{3}}{2}y + k - \frac{1}{2}l \\ \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l \end{pmatrix}. \quad (18.626)$$

Nous considérons  $(x, y) \in \text{Int}(K)$  tel que  $g(x, y) \in \text{Int}(K)$ , et nous allons trouver une contradiction. Le fait que  $(x, y) \in \text{Int}(K)$  signifie que  $(x, y)$  satisfait les inéquations (18.613) mais avec des inégalités strictes. Le fait que  $g(x, y) \in \text{Int}(K)$  nous donne trois inéquations de plus. Assez rapide calcul :

$$f_{AB}(g(x, y)) = -2x + k - 2l, \quad (18.627a)$$

$$f_{BC}(g(x, y)) = -x - \sqrt{3}y - k - l + 1, \quad (18.627b)$$

$$f_{AC}(g(x, y)) = \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l. \quad (18.627c)$$

Au final, notre point  $(x, y)$  doit satisfaire le système suivant :

$$\left\{ \begin{array}{l} x - \sqrt{3}y > 0 \end{array} \right. \quad (18.628a)$$

$$\left\{ \begin{array}{l} -x - \sqrt{3}y + 1 > 0 \end{array} \right. \quad (18.628b)$$

$$\left\{ \begin{array}{l} y > 0 \end{array} \right. \quad (18.628c)$$

$$\left\{ \begin{array}{l} -2x + k - 2l > 0 \end{array} \right. \quad (18.628d)$$

$$\left\{ \begin{array}{l} -x - \sqrt{3}y - k - l + 1 > 0 \end{array} \right. \quad (18.628e)$$

$$\left\{ \begin{array}{l} \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l > 0 \end{array} \right. \quad (18.628f)$$

Les inéquations 18.628c et 18.628a donnent déjà  $x > 0$ . De même avec 18.628b nous trouvons  $x < 1$ . Voilà déjà  $x \in ]0, 1[$  qui est directement visible sur le dessin du triangle  $K$ .

Puisque  $x > 0$ , l'inéquation (18.628d) donne

$$k - 2l > -2x + k - 2l > 0. \quad (18.629)$$

Donc  $k - 2l > 0$ .

Comme  $x < 1$  et  $y > 0$ , l'inéquation (18.628f) donne

$$\frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2}l > \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l > 0. \quad (18.630)$$

Donc  $\frac{\sqrt{3}}{2}(1 + l) > 0$ . Comme  $l$  est entier, cela donne  $l \geq 0$ .

Enfin, de (18.628e) nous tirons

$$-k + 1 > -x - \sqrt{3}y - k - l + 1 > 0, \quad (18.631)$$

Ce qui donne  $k < 1$  et donc  $k \leq 0$ .

En résumé nous avons trouvé trois inéquations pour  $k$  et  $l$  :

$$\left\{ \begin{array}{l} k \leq 0 \end{array} \right. \quad (18.632a)$$

$$\left\{ \begin{array}{l} l \geq 0 \end{array} \right. \quad (18.632b)$$

$$\left\{ \begin{array}{l} k - 2l > 0. \end{array} \right. \quad (18.632c)$$

Ce système est impossible.

Il n'existe donc pas de translation qui, appliquée à  $\text{Int}(L)$ , donne une intersection avec  $\text{Int}(K)$ .

(3) Pour  $L = R_0(\pi/3)^3 K$  hop.

(4) Pour  $L = R_0(\pi/3)^4 K$  hop.

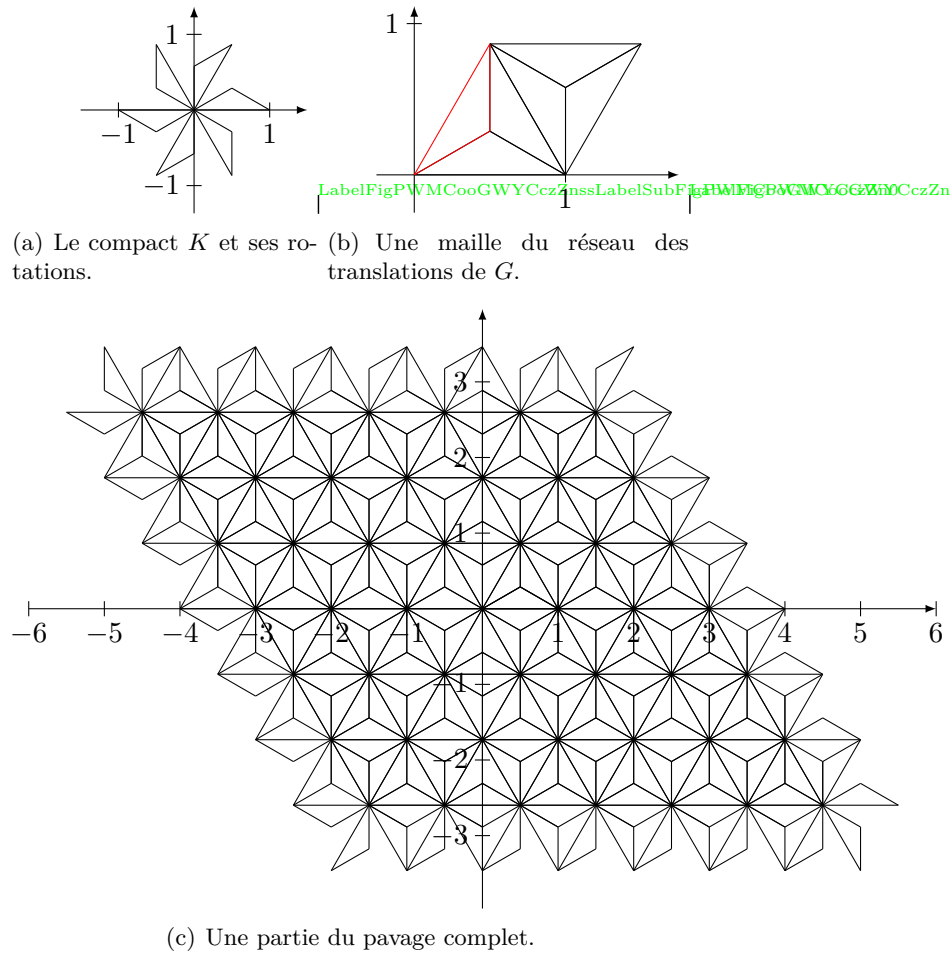


FIGURE 18.9: Illustrations pour le pavage du lemme 18.212.

(5) Pour  $L = R_0(\pi/3)^5 K$  hop.

Voilà. J'espère que toutes les idées sont en place, et que les parties manquantes sont seulement des vérifications qui se font mécaniquement, de la même manière<sup>69</sup>.  $\square$

### Lemme 18.213.

Le groupe

$$\text{gr} \left( \tau_1, \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(\pi/3) \right) \quad (18.633)$$

est un groupe de pavage.

D'après [618], la démonstration du lemme 18.213 demande d'utiliser le losange de sommets  $(0, 0)$ ,  $(1, 0)$ ,  $(\frac{1}{2}, \frac{\sqrt{3}}{6})$  et  $(\frac{1}{2}, -\frac{\sqrt{3}}{6})$ <sup>70</sup>.

### Lemme 18.214 ([1]).

Soient  $u_1, u_2 \in \mathbb{R}^2$  non colinéaires, ainsi que  $\alpha, \beta \in [0, 1]$ . Nous considérons  $v = \alpha u_1 + \beta u_2$ .

Nous supposons pour fixer les idées, que  $\|u_1\| \geq \|u_2\|$ . Alors

$$\min\{\|v\|, \|u_1 + u_2 - v\|\} \leq \|u_1\|. \quad (18.634)$$

Autrement dit, tout point intérieur d'un parallélogramme est plus proche d'un angle que la longueur du plus long côté.

69. Je n'ai pas vérifié. Faites-le et écrivez-moi pour me dire ce qu'il en est.

70. Je n'ai pas vérifié, mais à mon avis une preuve doit prendre les mêmes idées que celles du lemme 18.212.

*Démonstration.* Les points  $\alpha u_1 + \beta u_2$  ( $\alpha, \beta \in [0, 1]$ ) se divisent en deux parties : ceux avec  $0 \leq \alpha + \beta \leq 1$  et ceux avec  $1 \leq \alpha + \beta \leq 2$ .

Si  $\alpha + \beta \leq 1$  alors

$$\|\alpha u_1 + \beta u_2\| < \|\alpha u_1\| + \beta \|u_2\| = \alpha \|u_1\| + \beta \|u_2\| \leq (\alpha + \beta) \|u_1\| \leq \|u_1\|. \quad (18.635)$$

L'inégalité est stricte parce que  $u_1$  et  $u_2$  ne sont pas colinéaires.

Si au contraire  $\alpha + \beta \geq 1$  nous avons

$$\|u_1 + u_2 - \alpha u_1 - \beta u_2\| < \|(1 - \alpha)u_1\| + \|(1 - \beta)u_2\| < (2 - \alpha - \beta) \|u_1\| \leq \|u_1\|. \quad (18.636)$$

□

LEMooWK TGooQlfuxm

### Lemme 18.215.

Soient un ensemble  $E$ , et deux bijections  $r, s: E \rightarrow E$  ayant chacune un unique point fixe. Si elles commutent, alors leurs points fixes sont égaux.

*Démonstration.* Nous nommons  $a$  le point fixe de  $r$  et  $b$  celui de  $s$ . Pour tout  $x \in E$  nous avons  $(rs)(x) = (sr)(x)$ . En particulier pour  $x = s^{-1}(a)$ . D'une part

$$(rs)(x) = r(a) = a. \quad (18.637)$$

Et d'autre part,

$$(sr)(a) = (srs^{-1})(a) \quad (18.638)$$

Si nous imposons  $(srs^{-1})(a) = a$ , nous avons, en appliquant  $s^{-1}$  des deux côtés :  $(rs^{-1})(a) = s^{-1}(a)$ . Cela prouve que  $s^{-1}(a)$  est un point fixe de  $r$ . Donc  $s^{-1}(a) = a$ .

Nous en déduisons que  $a$  est un point fixe de  $s$  et donc que  $a = b$ .

□

LEMooDGSJooCiBhZz

### Lemme 18.216.

Soit un sous-groupe  $T$  de  $(\mathbb{R}^2, +)$  tel que  $\|v\| > \delta$  pour tout  $v \neq 0$  dans  $T$ .

Alors si  $u_1$  et  $u_2$  sont les plus petits éléments en norme de  $T$ , nous avons

$$T = \mathbb{Z}u_1 + \mathbb{Z}u_2. \quad (18.639)$$

*Démonstration.* Nous décomposons en plusieurs parties.

- (1)  $\mathbb{R}^+v \cap T = \mathbb{N}v_m$  Soit  $v \in T$ . L'ensemble  $\{\lambda \in \mathbb{R}^+ \text{ tel que } \lambda v \in T\}$  a un minimum parce que tous les éléments de  $T$  sont en norme plus grands que  $\delta > 0$ . Soit  $\lambda_m$  ce minimum et  $v_m = \lambda_m v$ .

Nous prétendons à présent que  $\mathbb{R}^+v \cap T = \mathbb{N}v_m$ . Nous ne faisons d'ailleurs pas que prétendre ; nous *prouvons*. En effet, soit  $\lambda v \in T$ . Nous devons prouver que  $\lambda = l\lambda_m$  pour un certain  $l \in \mathbb{N}$ . Soit  $k \in \mathbb{N}$  tel que  $k \leq \lambda < k + 1$ .

Nous avons

$$(k + 1)\lambda_m - \lambda \leq (k + 1)\lambda_m - k\lambda_m = \lambda_m. \quad (18.640)$$

Vu que  $T$  est un groupe pour l'addition, et que  $\lambda_m v \in T$  et  $\lambda v \in T$ , cela implique que  $((k + 1)\lambda_m - \lambda)v \in T$ . Mais

$$|((k + 1)\lambda_m - \lambda)v| \leq \lambda_m. \quad (18.641)$$

Vue la propriété de minimalité de  $\lambda_m$ , nous avons forcément

$$(k + 1)\lambda_m - \lambda = \lambda_m. \quad (18.642)$$

Cela prouve que  $\lambda = k\lambda_m$ .

Jusqu'ici nous avons prouvé que

$$\mathbb{R}^+v \cap T = \mathbb{N}v_m \quad (18.643)$$

pour un certain multiple  $v_m$  de  $v$ .

(2)  $\|u_1\| \leq \|v\|$  **pour tout**  $v$ 

Nous montrons à présent qu'il existe  $u_1 \in T$  tel que  $\|u_1\| \leq \|v\|$  pour tout  $v \in T$ . Si tel n'était pas le cas, il existerait une suite  $v_k \in T$  telle que  $\|v_{k+1}\| < \|v_k\|$ . Toute cette suite serait contenue dans la couronne (compacte) de rayons  $\delta$  et  $\|u_1\|$ . Quitte à prendre une sous-suite, nous pouvons supposer que  $(v_k)$  converge<sup>71</sup>. Cette suite serait de Cauchy et pour tout  $\epsilon$  (en particulier  $\epsilon < \delta$ ), il existerait  $p, q$  tels que  $\|v_p - v_q\| < \epsilon$ . Puisque  $T$  est un groupe pour l'addition, nous aurions  $v_p - v_q \in T$  avec  $\|v_p - v_q\| < \epsilon \leq \delta$ .

(3) **Première pause**

Si  $T$  est engendré seulement par  $u_1$ , nous avons fini. Autrement dit, si tout  $T$  est dans un sous-espace de dimension 1 de  $\mathbb{R}^2$ , nous avons terminé.

Dans la suite, nous supposons donc que  $T$  n'est pas contenu dans un sous-espace de dimension 1.

(4)  $T = \mathbb{Z}u_1 + \mathbb{Z}u_2$  Soient  $u_1$  et  $u_2$  les deux plus petits éléments de  $T$  en norme (peut-être ex-aequo). Ces deux éléments ne sont pas colinéaires, sinon leur différence serait plus petite. Ils forment donc une base de  $\mathbb{R}^2$ .

Soit  $v \in T$ . Comme  $\{u_1, u_2\}$  est une base de  $\mathbb{R}^2$ , il existe  $\alpha, \beta \in \mathbb{R}$  tels que

$$v = \alpha u_1 + \beta u_2. \quad (18.644)$$

Notre but est à présent de prouver que  $\alpha, \beta \in \mathbb{Z}$ .

Si ce n'était pas le cas, une simple translation nous mènerait dans les circonstances du lemme 18.214. Nous aurions alors que, soit  $\|v\|$ , soit  $\|u_1 + u_2 - v\|$  serait strictement plus petit que le plus grand entre  $\|u_1\|$  et  $\|u_2\|$ . Cela contredirait le fait que  $\|u_1\|$  et  $\|u_2\|$  étaient les deux plus petits.

□

PROPooPQYXooIDZiHy

**Proposition 18.217** ([1]).

Si  $G$  est un groupe de pavage<sup>72</sup> de  $\mathbb{R}^2$  et si  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  est une isométrie affine, alors le groupe

$$G' = f \circ G \circ f^{-1} = \{f \circ g \circ f^{-1} \text{ tel que } g \in G\} \quad (18.645)$$

est un groupe de pavage.

*Démonstration.* Soit un compact  $K$  tel que  $(G, K)$  soit un pavage. Nous notons  $K' = f(K)$ . Nous devons prouver deux choses :

- $f \circ G \circ f^{-1}$  est un groupe de déplacements ;
- $(G', K')$  est un pavage.

Ceci mène à prouver trois éléments.

(1)  $G'$  est constitué de déplacements Les éléments de  $G$  sont des isométries, ainsi que  $f$  et  $f^{-1}$ . Donc les éléments de  $G'$  sont des isométries.

Soit  $g \in G$ . Comme  $g$  est affine, il existe une décomposition  $g = \tau_v \circ g_0$  où  $g_0$  est linéaire. De même  $f = \tau_w \circ f_0$ . Les règles du produit et de l'inverse de la proposition 11.62(2)(3) nous indiquent que la partie linéaire de  $fgf^{-1}$  est  $f_0 g_0 f_0^{-1}$ .

En ce qui concerne le déterminant de  $f_0 g_0 f_0^{-1}$ , c'est la proposition 8.16 qui nous indique que

$$\det(f_0 g_0 f_0^{-1}) = \det(f_0) \det(g_0) \det(f_0^{-1}) = \det(f_0) \det(g_0) \det(f_0)^{-1} = \det(g_0) = 1. \quad (18.646)$$

71. Proposition 7.342.

72. Définition 18.207.

(2)  $\underline{G' \cdot K' = \mathbb{R}^2}$  Comme  $f$  est affine,  $K'$  est encore compact <sup>73</sup> Nous avons :

$$(f \circ G \circ f^{-1})(f(K)) = \bigcup_{g \in G} \bigcup_{k \in K} (f \circ g \circ f^{-1})(f(k)) \quad (18.647a)$$

$$= \bigcup_{g \in G} \bigcup_{k \in K} (f \circ g)(k) \quad (18.647b)$$

$$= \bigcup_{g \in G} f\left(\bigcup_{k \in K} g(k)\right) \quad (18.647c)$$

$$= \bigcup_{g \in G} (f \circ g)(K) \quad (18.647d)$$

$$= f\left(\bigcup_{g \in G} g(K)\right) \quad (18.647e)$$

$$= f(G \cdot K) \quad (18.647f)$$

$$= \mathbb{R}^2. \quad (18.647g)$$

Pour la dernière égalité, nous avons utilisé le fait que  $f$  est bijective et que  $G \cdot K = \mathbb{R}^2$ .

(3) **L'autre condition** Deux éléments de  $G'$  s'écrivent  $f g_1 f^{-1}$  et  $f g_2 f^{-1}$ . Nous les supposons tels que

$$(f g_1 f^{-1}) \cdot \text{Int}(K') \cap (f g_2 f^{-1}) \cdot \text{Int}(K') \neq \emptyset. \quad (18.648)$$

Nous avons :

$$(f g_1 f^{-1}) \cdot \text{Int}(K') \cap (f g_2 f^{-1}) \cdot \text{Int}(K') = f(g_1 \cdot \text{Int}(K)) \cap f(g_2 \cdot \text{Int}(K)) \quad (18.649a)$$

$$= f\left(\underbrace{g_1 \cdot \text{Int}(K) \cap g_2 \cdot \text{Int}(K)}_{\neq \emptyset}\right) \quad (18.649b)$$

$$\neq \emptyset. \quad (18.649c)$$

Justifications :

- Égalité (18.649a) parce qu'un peu de topologie nous enseigne que  $f^{-1}(\text{Int}(K')) = \text{Int}(f^{-1}(K')) = \text{Int}(K)$  parce que  $f$  est affine.
- Égalité (18.649b) parce que,  $f$  étant bijective,  $f(A) \cap f(B) = f(A \cap B)$ ;

□

THOooUPHQooYfeHAy

**Théorème 18.218** ([486, 618]).

Nous notons  $\tau_v$  la translation de vecteur  $v$ ,  $r_{A,\theta}$  la rotation de centre  $A$  et d'angle  $\theta$  ainsi que  $\tau_i = \tau_{e_i}$ .

Un groupe  $G$  est un groupe de pavage de  $\mathbb{R}^2$  si et seulement si il existe une bijection affine  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  telle que  $f \circ G \circ f^{-1}$  est un groupe de la liste suivante :

$$(1) \text{ gr } (\tau_1, \tau_2)$$

$$(2) \text{ gr } (\tau_1, \tau_2, R_0(\pi))$$

$$(3) \text{ gr } (\tau_1, \tau_{(\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(2\pi/3))$$

$$(4) \text{ gr } (\tau_1, \tau_2, R_0(\pi/2))$$

$$(5) \text{ gr } (\tau_1, \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(\pi/3)).$$

*Démonstration.* Les lemmes 18.209, 18.210, 18.211, 18.212, et 18.213 montrent que les groupes listés sont des groupes de pavage. La proposition 18.217 nous montre alors que si  $H = fGf^{-1}$  est dans la liste, alors  $G$  est un groupe de pavage. Il nous reste à montrer que si  $G$  est un groupe de pavage, alors il existe une application affine  $\alpha$  telle que  $\alpha G \alpha^{-1}$  est un groupe de la liste.

Soit  $(G, K)$  un pavage de  $\mathbb{R}^2$ . Nous notons  $T$  l'ensemble des translations dans  $G$ , plus précisément,

$$T = \{v \in \mathbb{R}^2 \text{ tel que } \tau_v \in G\}. \quad (18.650)$$

73. Il suffit de prouver que  $f(K)$  est fermé et borné par le théorème de Borel-Lebesgue 9.24.

- (1) **Une borne pour  $T$**  Nous prouvons qu'il existe  $\delta > 0$  tel que  $\|v\| > \delta$  pour tout  $v \neq 0 \in T$ . En effet, soit  $m \in \text{Int}(K)$  ainsi que  $r$  tel que  $B(m, r) \subset \text{Int}(K)$ . Alors si  $v \in T$  est tel que  $\|v\| < r$  nous avons  $\tau_v(m) \in \text{Int}(K)$ , ce qui donnerait

$$m \in \tau_v(\text{Int}(K)) \cap \text{Int}(K) \quad (18.651)$$

par hypothèse, cette intersection est non vide seulement si  $v = 0$ .

Donc il existe  $\delta$  tel que  $\|v\| \geq \delta$  pour tout  $v \in T$ .

- (2) **Utilisation du lemme**

La partie  $T$  est donc dans la position du lemme 18.216 et nous avons

$$T = \mathbb{Z}u_1 + \mathbb{Z}u_2 \quad (18.652)$$

pour les vecteurs  $u_1$  et  $u_2$  les plus petits en norme de  $T$ .

En réalité, il se peut que  $T$  soit plus petit que ça, parce que  $G$  peut par exemple ne contenir aucune translation. Nous avons trois possibilités :

- $T = \{0\}$ ,
- $T = \mathbb{Z}u$  pour un certain  $u \neq 0$  dans  $\mathbb{R}^2$ ,
- $T = \mathbb{Z}u_1 + \mathbb{Z}u_2$  pour certains  $u_1, u_2 \in \mathbb{R}^2 \setminus \{0\}$ .

- (3) **Translation** Si  $r, s \in G$ , alors l'élément  $rsr^{-1}s^{-1}$  est une translation. En effet, puisque les éléments de  $G$  sont des déplacements, ce sont des applications affines et donc il existe des applications linéaires  $A_r, A_s$  et des translations  $\tau_r, \tau_s$  telles que  $r = A_r \circ \tau_r$  et  $s = A_s \circ \tau_s$ . La proposition 11.62 nous donne les inverses. Nous avons

$$rsr^{-1}s^{-1} = (A_r \circ \tau_r)(A_s \circ \tau_s)(A_r^{-1} \circ \tau_{-A_r v_r})(A_s^{-1} \circ \tau_{-A_s v_s}). \quad (18.653)$$

La partie linéaire de cela est

$$A_r \circ A_s \circ A_r^{-1} \circ A_s^{-1}. \quad (18.654)$$

C'est donc une composée de rotations centrées en  $(0, 0)$ . Mais ces rotations forment un groupe abélien (proposition 18.128). Donc nous pouvons écrire

$$A_r \circ A_s \circ A_r^{-1} \circ A_s^{-1} = A_r \circ A_r^{-1} \circ A_s \circ A_s^{-1} = \text{Id}. \quad (18.655)$$

Tout ceci pour dire que dès que  $r, s \in G$ , l'élément  $rsr^{-1}s^{-1}$  est une translation.

- (4) **Les parties linéaires**[1] Nous savons de l'exemple 11.19 que les éléments de  $G$  s'écrivent sous la forme  $f = \tau_v \circ \alpha$  où  $v \in \mathbb{R}^2$  et  $\alpha: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  est linéaire.

De plus,  $f$  étant une isométrie de  $(\mathbb{R}^2, d)$ , l'application  $\alpha$  est une isométrie de  $(\mathbb{R}^2, \|\cdot\|)$ . Puisque  $\alpha$  est une isométrie,  $\det(\alpha) = \pm 1$ . Mais les déplacements conservent l'orientation ; donc  $\alpha$  doit conserver l'orientation, et la proposition 8.33 nous dit que  $\det(\alpha) > 0$ . Donc

$$\det(\alpha) = 1. \quad (18.656)$$

Le théorème 18.96 dit que  $\alpha$  est la composition d'un nombre pair de réflexions. Mais comme il y en a au plus trois (théorème 18.188), l'application  $\alpha$  est composée de zéro ou deux réflexions.

Donc les parties linéaires des éléments de  $G$  sont des rotations.

- (5) **Les autres**

Les parties linéaires des éléments de  $G$  sont des rotations. Mais les éléments de  $G$  eux-mêmes ne sont pas tellement mystérieux. Puisque ce sont des isométries de  $(\mathbb{R}^2, d)$ , elles sont composées de 0, 1, 2 ou 3 réflexions.

Mais ce sont des déplacements, donc ils préservent l'orientation et le théorème 18.96 dit qu'ils sont des composées de zéro ou deux réflexions (nombre pair). Ce sont donc des rotations.

- (6) **Hein ?** Les éléments linéaires de  $G$  sont des rotations. Et les autres aussi ? Les linéaires sont des rotations autour de  $(0, 0)$  ; les autres sont des rotations autour de points autres que  $(0, 0)$ .

C'est pourquoi dans la suite, nous précisons « rotation linéaire » pour une rotation autour de  $(0, 0)$  et nous dirons « rotation » pour une rotation en général. Dans le contexte affine, il faut toujours faire attention à ça : une rotation peut très bien n'être pas linéaire <sup>74</sup>.

- (7) **Les rotations linéaires stabilisent  $T$**  Nous prouvons maintenant que les rotations linéaires de  $G$  stabilisent  $T$ , c'est-à-dire que si  $v \in T$  et si  $\alpha$  est une rotation linéaire de  $G$ , alors  $\alpha(v) \in T$ . La transformation  $\alpha\tau_v\alpha^{-1}$  est dans  $G$ . Mais pour tout  $x \in \mathbb{R}^2$  nous avons

$$(\alpha\tau_v\alpha^{-1})(x) = \alpha(\alpha^{-1}(x) + v) = x + \alpha(v) = \tau_{\alpha(v)}(x). \quad (18.657)$$

Donc  $\alpha\tau_v\alpha^{-1} = \tau_{\alpha(v)}$  et  $\alpha(v) \in T$ .

- (8) **Exclusion de  $T = \{0\}$**  Le fait que  $T = \{0\}$  ne signifie pas que tous les éléments de  $G$  sont des rotations ; il peut encore y avoir des composées de rotations et de translations  $A \circ \tau$ . Cela étant dit, si  $T = \{0\}$ , il n'en reste pas moins que  $rsr^{-1}s^{-1}$  est une translation, c'est-à-dire est égal à Id. Mais  $rsr^{-1}s^{-1} = e$  implique  $rs = sr$ .

Donc  $G$  est abélien. Les éléments de  $G$  sont donc des rotations qui commutent deux à deux. Puisqu'une rotation a son centre comme unique point fixe, le lemme 18.215 nous dit que tous les éléments de  $G$  sont des rotations de même centre.

Soit  $c$  le centre commun de tous les éléments de  $G$ . Vu que  $K$  est compact dans  $\mathbb{R}^2$ , il existe  $r > 0$  tel que  $K \subset B(c, r)$ . Puisque  $G$  stabilise toutes les boules centrées en  $c$ , nous avons

$$G \cdot K \subset B(c, r). \quad (18.658)$$

Donc nous n'avons pas un recouvrement de  $\mathbb{R}^2$ . Le cas  $T = \{0\}$  est exclu.

- (9) **Exclusion de  $T = \mathbb{Z}u$**  Nous supposons à présent que  $T = \mathbb{Z}u$  pour un certain  $u \in \mathbb{R}^2$ .

- (9a)  **$r_0 = \pm \text{Id}$**  Nous savons que tous les éléments de  $G$  sont des rotations ; soit un élément  $r$  de  $G$ . La proposition 18.127(3) nous indique qu'il existe un point  $a \in \mathbb{R}^2$  ainsi qu'une rotation linéaire  $r_0$  telle que  $r = \tau_a^{-1}r_0\tau_a$ . Nous allons prouver que  $r_0$  est  $\pm \text{Id}$ . D'abord,

$$r\tau_u r^{-1} = \tau_a^{-1}r_0\tau_a\tau_u\tau_a^{-1}r_0^{-1}\tau_a = \tau_a^{-1}r_0\tau_u r_0^{-1}\tau_a. \quad (18.659)$$

Ensuite, nous appliquons cela à  $x \in \mathbb{R}^2$  :

$$(\tau_a^{-1}r_0\tau_u r_0^{-1}\tau_a)(x) = (\tau_a^{-1}r_0\tau_u)(r_0^{-1}(x + a)) \quad (18.660a)$$

$$= (\tau_a^{-1}r_0)(r_0^{-1}(x + a) + u) \quad (18.660b)$$

$$= \tau_a^{-1}(x + a + r_0(u)) \quad (18.660c)$$

$$= x + r_0(u). \quad (18.660d)$$

Donc  $r \circ \tau_u \circ r^{-1} = \tau_{r_0(u)}$ , ce qui prouve que  $r_0(u) \in T$ . Comme  $\|r_0(u)\| = \|u\|$  nous avons forcément  $r_0(u) = \pm u$ .

Si  $r_0(u) = u$ , alors  $r_0 = \text{Id}$  (parce que  $r_0$  est une rotation fixant plus qu'un seul point). Dans ce cas,  $r = \text{Id}$ .

Si au contraire  $r_0(u) = -u$ , alors  $r_0 = -\text{Id}$ .

- (9b) **Forme générale** Donc si  $r$  est un élément non trivial de  $G$  nous avons  $r = \tau_a^{-1} \circ (-\text{Id}) \circ \tau_a$ , et alors

$$(\tau_a^{-1} \circ (-\text{Id}) \circ \tau_a)(x) = \tau_a^{-1}(-\text{Id})(x + a) = \tau_a^{-1}(-x - a) = -x - 2a. \quad (18.661)$$

Donc pour tout  $r \in G$ , il existe  $a \in \mathbb{R}^2$  tel que

$$r(x) = -x - 2a. \quad (18.662)$$

Pour information, le centre de cette rotation est  $-a$  (c'est le seul point fixe).

74. Lorsque, ailleurs dans le Frido, nous disons « rotation », souvent nous pensons « rotation linéaire ». Gardez cependant à l'esprit qu'une rotation peut très bien être centrée ailleurs qu'en l'origine, et soyez toujours capable de préciser le cas échéant.

(9c) **Les centres sont alignés** Soit  $r$  une rotation de centre  $-a$  et  $s$  de centre  $-b$ . Alors

$$(rs)(x) = r(-x - 2b) = x + 2b - 2a = x + 2(b - a). \quad (18.663)$$

Donc  $rs = \tau_{2(b-a)}$ .

Cela prouve que  $2(b - a) \in \mathbb{Z}u$ .

(9d) **Une bande** Soit la droite  $D = \mathbb{R}u$ . Nous considérons la bande

$$B_r = \{x \in \mathbb{R}^2 \text{ tel que } d(x, D) < r\}. \quad (18.664)$$

(9e) **Une inclusion** Nous prouvons à présent que pour tout  $r$ , nous avons

$$G \cdot B_r \subset B_r. \quad (18.665)$$

Nous savons qu'un élément général de  $G$  est une rotation centrée en un point de  $D$ , et que l'action d'une telle rotation est donnée par (18.662). Nous avons

$$\begin{aligned} d(r(x), D) &= d(x + 2a, D) && \text{EQooNCQGooGGAhCR} && (18.666a) \\ &= d(\tau_{2a}^{-1}(x + 2a), \tau_{2a}^{-1}(D)) && \text{EQooNFJNooSAYtBD} && (18.666b) \\ &= d(x, D). && \text{SUBEQooMGT MooAlRwWT} && (18.666c) \end{aligned}$$

Justifications :

- Pour (18.666a), nous avons  $D = -D$  et  $d(x, y) = d(-x, -y)$ .
- Pour (18.666b), invariance par translation de la distance dans  $\mathbb{R}^2$ .
- Pour (18.666c), les éléments de  $D$  sont les multiples de  $a$ ; donc cette droite est invariante par cette translation.

Bref,  $d(x, D) = d(r(x), D)$ . Donc, pour tout  $r > 0$ <sup>75</sup> et pour tout  $g \in G$ , si  $x \in B_r$ , alors  $g(x) \in B_r$ .

(9f) **Exclusion**

Comme  $K$  est compact et que la fonction  $x \mapsto d(x, D)$  est continue, il existe  $r > 0$  tel que  $K \subset B_r$ . Avec ça,  $G \cdot K \subset G \cdot B_r \subset B_r$ . Donc  $G \cdot K$  ne recouvre pas tout  $\mathbb{R}^2$  et  $G$  n'est pas un groupe de pavage.

Et nous voilà avec seulement  $T = \mathbb{Z}u_1 + \mathbb{Z}u_2$  en lice.

(1) **Pause : quelques parties de  $G$  à ne pas confondre** Il convient de ne pas se perdre entre différentes parties de  $G$ . Je vous laisse méditer quelques temps sur la liste suivante :

- (1a)  $G$  est le groupe que nous cherchons à déterminer ;
- (1b)  $T$  est le groupe des translations de  $G$  ;
- (1c) le groupe des rotations linéaires dans  $G$  ;
- (1d) l'ensemble des  $\tau_A^{-1} \circ r \circ \tau_A$  où  $r$  est une rotation de  $G$  centrée en  $A$ .<sup>ITEMooFWWMooNzLUGy</sup>
- (1e)  $L$  est l'ensemble des parties linéaires des éléments de  $G$ .<sup>ITEMooIEJZooNaSKpc</sup>

En particulier les deux derniers points ne sont pas les mêmes.

Dans le cas (1d), il s'agit de dire que  $r$  est une rotation centrée en  $A \in \mathbb{R}^2$  et écrire  $r = \tau_A \circ r_0 \circ \tau_A^{-1}$  (proposition 18.127(3)) et considérer  $r_0$ . Dans ce cas,  $r_0$  est une rotation, mais elle n'est pas ce que nous appelons la « partie linéaire » de  $r$ . Il n'y a pas de garantie que cela forme un groupe.

Si  $r$  est une rotation dans  $G$ , dans le cas (1e) il s'agit de décomposer  $r = \tau_v \circ \alpha$  (lemme 11.14) et considérer  $\alpha$ . Dans ce cas,  $\alpha$  est linéaire, mais il n'y a pas de garantie que  $\alpha$  soit une rotation.

(2)  **$L$  est un groupe** En trois conditions.

<sup>75</sup>. Remarquez la notation malheureuse pour  $r$  qui est maintenant une distance alors que trois mots plus tôt, c'était un élément de  $G$ .



- Si  $\alpha, \beta \in L$ , il existe  $v, w \in \mathbb{R}^2$  tels que  $\tau_v \circ \alpha \in G$  et  $\tau_w \circ \beta \in G$ . La loi de produit 11.62(2) dit que  $\tau_v \circ \alpha \circ \tau_w \circ \beta = \tau_{\alpha(w)+v} \circ \alpha\beta$ . Donc  $\alpha\beta$  est la partie linéaire d'un élément de  $G$ .
- De la même façon, en utilisant l'inverse 11.62(3),  $(\tau_v \circ \alpha)^{-1} = \tau_{-\alpha^{-1}(v)} \circ \alpha^{-1}$ . Donc  $\alpha^{-1}$  est la partie linéaire d'un élément de  $G$ .
- Et enfin  $\text{Id} = \tau_0 \circ \text{Id}$ . Donc l'identité est dans  $L$ .

Ok :  $L$  est un groupe.

- (3) **Précision** L'ensemble  $L$  est un groupe, certes. Mais rien ne dit que  $L$  soit un sous-groupe de  $G$ .
- (4)  **$L$  préserve le réseau** Soit  $\alpha \in L$ . Il existe  $v \in \mathbb{R}^2$  tel que  $g = \tau_v \circ \alpha \in G$ . Soit  $u \in T$ . Nous allons montrer que  $\alpha(u) \in T$ . Vu que  $g$  et  $\tau_u$  sont dans  $G$ , l'élément  $g\tau_u g^{-1}$  est également dans  $G$ . Nous l'appliquons à  $x \in \mathbb{R}^2$  :

$$(\tau_v \alpha \tau_u \alpha^{-1} \tau_v^{-1})(x) = (\tau_v \alpha \tau_u \alpha^{-1})(x - v) \quad (18.667a)$$

$$= (\tau_v \alpha)(\alpha^{-1}(x - v) + u) \quad (18.667b)$$

$$= \tau_v(x - v + \alpha(u)) \quad (18.667c)$$

$$= x + \alpha(u) \quad (18.667d)$$

$$= \tau_{\alpha(u)}(x). \quad (18.667e)$$

Donc  $g\tau_u g^{-1} = \tau_{\alpha(u)} \in G$ .

- (5) **Question de trace** Soit  $\alpha \in L$  ; dans la base  $\{u_1, u_2\}$  la matrice de  $\alpha$  est

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (18.668)$$

avec  $a, b, c, d \in \mathbb{Z}$ . La trace de cette matrice est  $a + d \in \mathbb{Z}$ . Dans la base canonique de  $\mathbb{R}^2$  par contre, la proposition 18.138 nous dit qu'il existe  $\theta \in [0, 2\pi[$  tel que la matrice de  $\alpha$  soit

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (18.669)$$

La trace est  $2\cos(\theta)$ . La trace est invariante par changement de base<sup>76</sup>, donc  $2\cos(\theta) = a + d \in \mathbb{Z}$ . Les possibilités pour  $\cos(\theta)$  sont donc  $-1, -1/2, 0, 1/2$  et  $1$ .

- (6) **Les angles possibles** Nous savons que  $\cos(\frac{\pi}{2}) = 0$  et  $\cos(\pi/3) = 1/2$  (proposition 18.21(6) et lemme 18.28). La proposition 18.75 nous dit alors que, dans notre cas, les valeurs possibles pour  $\theta$  dans  $[0, 2\pi[$  sont

$$\{0, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \pi, \frac{4\pi}{3}, \frac{3\pi}{2}, \frac{5\pi}{3}\}. \quad (18.670)$$

Donc les rotations possibles dans  $L$  sont les rotations de ces angles.

Nous devons trouver quels sont les groupes qui peuvent être formés seulement avec ces éléments.

- (7) **Quelques combinaisons impossibles**

Puisque  $L$  est un groupe, il y a des combinaisons impossibles. Par exemple si  $R_0(\pi/3)$  et  $R_0(\pi/2)$  sont dans  $L$ , alors la composée<sup>77</sup>  $R_0(\pi/2)R_0(\pi/3) = R_0(5\pi/6)$  est également dans  $L$ . Mais comme  $5\pi/6$  n'est pas dans la liste (18.670),  $R_0(5\pi/6)$  n'est pas dans  $L$ .

En raisonnant de la sorte, nous voyons que si  $R_0(\pi/2) \in L$ , alors  $L = \text{gr}(R_0(\pi/2))$ .

- (8) **La liste** Plus généralement, les possibilités pour  $L$  sont

- $\{\text{Id}\}$
- $\text{gr}(R_0(\pi/2))$
- $\text{gr}(R_0(\pi/3))$

<sup>76</sup>. Proposition 8.203.

<sup>77</sup>. Proposition 18.139 pour l'addition des angles.

- $\text{gr}(R_0(\pi))$
- $\text{gr}(R_0(2\pi/3))$

Une justification plus courte pour cette liste est d'invoquer le théorème 18.199(2) qui dit que  $L$  étant un sous-groupe fini des isométries de  $(\mathbb{R}^2, d)$ , il est cyclique et donc monogène. Notons pour cela que  $\text{gr}(R_0(4\pi/3)) = \text{gr}(R_0(2\pi/3))$  parce que  $R_0(4\pi/3) = R_0(2\pi/3)^{-1}$ . De la même manière  $R_0(3\pi/2) = R_0(\pi/2)^{-1}$  et  $R_0(5\pi/3) = R_0(\pi/3)^{-1}$ .

(9) **Le cas un peu générique**

Nous supposons que  $L = \text{gr}(R_0(\theta))$  pour un certain  $\theta$ . Nous allons voir qu'à part dans les cas  $\theta = 0$  et  $\theta = \pi$ , il est possible de trouver une application affine  $\alpha$  telle que le groupe  $\alpha G \alpha^{-1}$  soit alors dans la liste.

Nous considérons un élément de  $G$  de la forme  $\tau_{v_0} \circ R_0(\theta)$ . Pour être bien clair, il n'est absolument pas garanti que  $v_0$  soit dans  $T$ . Nous allons chercher un élément  $w_0 \in \mathbb{R}^2$  tel que le groupe de pavage<sup>78</sup>  $G' = \tau_{w_0} G \tau_{w_0}^{-1}$  contienne  $R_0(\theta)$ . Le groupe  $G'$  contient l'élément  $g = \tau_{w_0} \tau_{v_0} R_0(\theta) \tau_{w_0}^{-1}$ ; nous l'appliquons à  $x \in \mathbb{R}^2$  :

$$g(x) = R_0(\theta)x - R_0(\theta)w_0 + w_0 + v_0. \quad (18.671)$$

Nous avons  $g(x) = R_0(\theta)x$  lorsque

$$w_0 = (R_0(\theta) - \mathbb{1})^{-1}v_0. \quad (18.672)$$

Voilà pourquoi le cas  $\theta = 0$  sera traité à part : dans le cas  $\theta = 0$ , l'opérateur  $R_0(\theta) - \mathbb{1}$  n'est pas inversible. Dans les autres cas, nous avons un groupe  $G' = \tau_{w_0} G \tau_{w_0}^{-1}$  qui contient  $R_0(\theta)$ . Puisqu'un élément général de  $G$  est de la forme  $\tau_v \circ R_0(\theta)^k$ , un élément général de  $G'$  est de la forme

$$\tau_{w_0} \circ \tau_v \circ R_0(\theta)^k \circ \tau_{w_0}^{-1} = \tau_{w_0+v-R_0(\theta)^k w} \circ R_0(\theta)^k. \quad (18.673)$$

Donc tous les éléments de  $G'$  sont encore de la forme

$$\tau_v \circ R_0(\theta)^k. \quad (18.674)$$

Mais dans  $G'$  nous avons une information capitale :  $R_0(\theta)^k$  lui-même est dans  $G'$ . Donc si  $\tau_v \circ R_0(\theta)^k \in G'$ , alors  $\tau_v \in G'$ .

Vu que  $G'$  est encore un groupe de pavage, tout ce qui a été dit précédemment tient, et le groupe des translations dans  $G'$  est un réseau  $T' = \mathbb{Z}u_1 + \mathbb{Z}u_2$  où  $u_1$  et  $u_2$  peuvent être choisis arbitrairement parmi les deux plus petits vecteurs non colinéaires de  $T$ .

Tous les éléments de  $G'$  sont de la forme  $\tau_v \circ R_0(\theta)^k$  avec  $v \in T'$ . Donc

$$G' = \text{gr}(\tau_{u_1}, \tau_{u_2}, R_0(\theta)). \quad (18.675)$$

Nous savons que  $R_0(\theta)$  fixe  $T'$ . Or l'élément  $R_0(\theta)u_1$  a la même norme que  $u_1$ ; donc nous pouvons choisir  $u_2 = R_0(\theta)u_1$  pour peu que  $R_0(\theta)u_1$  soit non colinéaire à  $u_1$ . Et c'est ici que nous laissons le cas  $\theta = \pi$  de côté.

Nous écrivons donc sans vergogne que

$$G' = \text{gr}(\tau_{u_1}, \tau_{R_0(\theta)u_1}, R_0(\theta)). \quad (18.676)$$

Nous allons maintenant nous occuper de  $u_1$ . Pour cela nous considérons une rotation suivie d'une homothétie  $\alpha$  telle que  $\alpha(u_1) = e_1$ . Une telle opération  $\alpha$  d'une part, commute avec  $R_0(\theta)$  et d'autre part, réalise  $\alpha \circ \tau_v \circ \alpha^{-1} = \tau_{\alpha(v)}$ . Donc le groupe  $G'' = \alpha G' \alpha^{-1}$  est, par le lemme 1.426,

$$G'' = \text{gr}(\tau_{e_1}, \tau_{\alpha R_0(\theta)u_1}, R_0(\theta)). \quad (18.677)$$

Cela s'écrit aussi bien

$$G'' = \text{gr}(\tau_{e_1}, \tau_{R_0(\theta)e_1}, R_0(\theta)). \quad (18.678)$$

Et voilà, ce groupe  $G''$  est un de ceux de la liste.

---

78. Proposition 18.217.

- (10) **Le cas  $L = \{\text{Id}\}$**  Dans ce cas, tous les éléments de  $G$  sont de la forme  $\tau_v$  avec  $v \in T$ . Nous considérons l'application linéaire  $\alpha: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  telle que  $\alpha(u_1) = e_1$  et  $\alpha(u_2) = e_2$ . Nous avons

$$(\alpha \circ \tau_{u_i} \circ \alpha^{-1})(x) = (\alpha \tau_{u_i}) \alpha^{-1}(x) = \alpha(\alpha^{-1}(x) + u_i) = x + \alpha(u_i) = \tau_i(x). \quad (18.679)$$

Donc  $\alpha \circ \tau_{u_i} \circ \alpha^{-1} = \tau_i$ . De la même façon, si  $a \in \mathbb{Z}$  nous avons  $\alpha \tau_{au_i} \alpha^{-1} = \tau_{ae_i}$ . Et avec tout ça, si  $v \in T$ , alors  $v = au_1 + bu_2$  et nous avons

$$\alpha \tau_v \alpha^{-1} = \alpha \tau_{au_1} \tau_{bu_2} \alpha^{-1} = \alpha \tau_{au_1} \alpha^{-1} \alpha \tau_{bu_2} \alpha^{-1} = \tau_{ae_1} \tau_{be_2}. \quad (18.680)$$

Nous avons donc

$$\alpha G \alpha^{-1} = \text{gr}(\tau_1, \tau_2). \quad (18.681)$$

Cela est un des groupes de la liste.

Notez qu'à la place de ces calculs, nous pouvions aussi invoquer la proposition 1.426.

- (11) **Le cas  $L = \{-\text{Id}\}$**  Dans le cas  $\theta = \pi$ , nous pouvons aller jusqu'à (18.675) et écrire

$$G' = \text{gr}(\tau_{u_1}, \tau_{u_2}, R_0(\pi)). \quad (18.682)$$

Puisque  $R_0(\pi) = -\text{Id}$  commute avec toutes les applications linéaires et que  $u_1$  et  $u_2$  ne sont pas colinéaires, nous pouvons considérer une application linéaire  $\alpha: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  telle que  $\alpha(u_1) = e_1$ ,  $\alpha(u_2) = e_2$ . Nous avons alors

$$G'' = \alpha G' \alpha^{-1} = \text{gr}(\tau_{e_1}, \tau_{e_2}, -\text{Id}), \quad (18.683)$$

qui est encore dans la liste.

□

## 18.12 Un peu de structure de $O(n)$

### 18.12.1 Valeurs propres dans $O(n)$

**Proposition 18.219** ([619]).

Soit une matrice  $A \in O(n)$ . Si  $\lambda \in \mathbb{C}$  est une valeur propre de  $A$ , alors  $\bar{\lambda}$  est également une valeur propre de  $A$ , et de plus  $|\lambda| = 1$ .

*Démonstration.* Dire que  $\lambda \in \mathbb{C}$  est une valeur propre de  $A$  signifie qu'il existe  $x \in \mathbb{C}^n$  (non nul) tel que  $Ax = \lambda x$ . Comme les éléments de la matrice  $A$  sont réels,

$$A\bar{x} = \bar{A}\bar{x} = \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x}. \quad (18.684)$$

Donc  $\bar{\lambda}$  est une valeur propre de  $A$  pour le vecteur propre  $\bar{x}$ .

Soit  $\lambda$  une valeur propre de  $A$  de vecteur propre  $x$ . Alors nous avons d'une part

$$\langle \overline{Ax}, Ax \rangle = \langle A^t A \bar{x}, x \rangle = \langle x, \bar{x} \rangle, \quad (18.685)$$

et d'autre part

$$\langle \overline{Ax}, Ax \rangle = \langle \bar{\lambda}\bar{x}, \lambda x \rangle = |\lambda|^2 \langle \bar{x}, x \rangle. \quad (18.686)$$

Puisque  $x \neq 0$  nous avons aussi  $\langle \bar{x}, x \rangle \neq 0$ . Par conséquent  $|\lambda|^2 = 1$  et  $|\lambda| = 1$ . □

**Lemme 18.220** ([619]).

Soient un espace vectoriel euclidien  $E$  de dimension finie et une isométrie  $f$  de  $E$ . Soit  $F$  un sous-espace de  $E$  stable par  $f$ . Alors  $F^\perp$  est stable par  $f$ .

*Démonstration.* La restriction  $f_F: F \rightarrow F$  est encore une isométrie ; elle est donc inversible : pour tout  $y \in F$ , il existe  $x \in F$  tel que  $y = f(x)$ . Soit  $a \in F^\perp$  ; nous montrons que  $f(a) \in F^\perp$ . Soit donc  $y \in F$  et calculons :

$$\langle y, f(a) \rangle = \langle f(x), f(a) \rangle = \langle x, a \rangle = 0 \quad (18.687)$$

parce que  $x \in F$  et  $a \in F^\perp$ . □

**Proposition 18.221** ([619]).

Soit une isométrie  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ .

- (1) L'application linéaire  $f$  possède au moins une valeur propre réelle qui vaut  $\pm 1$ .
- (2) Il existe une base orthonormée de  $\mathbb{R}^3$  dans laquelle la matrice de  $f$  est de la forme

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos(\theta) & -\epsilon \sin(\theta) \\ 0 & \sin(\theta) & \epsilon \cos(\theta) \end{pmatrix} \quad (18.688)$$

avec  $\epsilon, \lambda = \pm 1$  et  $\theta \in [0, 2\pi[$ .

*Démonstration.* Le polynôme caractéristique de  $f$ , donné par  $\det(f - \lambda \text{Id})$ , est à coefficients réels et de degré 3. Il possède donc au moins une solution réelle, par le corolaire 12.93. Soit donc une valeur propre réelle  $\lambda$  de  $\chi_f$ ; par le lemme 18.219 nous avons  $\lambda = \pm 1$ . Soit  $u_1$  le vecteur propre correspondant. Nous notons  $F$  l'espace engendré par  $u_1$ .

Nous avons  $f(F) = F$  et donc  $f(F^\perp) = F^\perp$  par le lemme 18.220. Soit une base orthonormée  $\{u_2, u_3\}$  de  $F^\perp$  et la matrice  $B$  de la restriction  $f_p$  à  $F^\perp$ . Comme l'application  $f_p$  est une isométrie de  $F^\perp$ , la matrice  $B$  est, par le lemme 18.134, de la forme

$$B = \begin{pmatrix} \cos(\theta) & -\epsilon \sin(\theta) \\ \sin(\theta) & \epsilon \cos(\theta) \end{pmatrix} \quad (18.689)$$

pour un certain  $\theta \in [0, 2\pi[$  et  $\epsilon = \pm 1$ .

Dans la base  $\{u_1, u_2, u_3\}$  de  $\mathbb{R}^3$ , la matrice de  $f$  est alors

$$\begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix}, \quad (18.690)$$

comme annoncé. □

Pour classifier les isométries de  $\mathbb{R}^3$ , nous pouvons nous baser sur les possibilités de la matrice donnée dans le lemme 18.221. Il y a essentiellement quatre possibilités suivant les valeurs de  $\lambda = \pm 1$  et  $\epsilon = \pm 1$ .

- (1) Si  $\epsilon = \lambda = 1$  Alors la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \quad (18.691)$$

et l'isométrie correspondante est la rotation d'angle  $-\theta$  autour de la droite de  $u_1$ .

- (2) Si  $\epsilon = \lambda = -1$  Alors la matrice est

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{pmatrix} \quad (18.692)$$

Cette application est plus subtile, parce que même dans le plan  $\text{Span}(u_2, u_3)$ , ce n'est pas une rotation. Nous allons montrer qu'il s'agit d'une réflexion autour de la droite d'angle  $\theta/2$  dans le plan  $\text{Span}(u_2, u_3)$ . Nous nommons  $D$  cette droite. Dans la base  $(u_1, u_2, u_3)$ , cette droite est donnée par <sup>79</sup>

$$t \mapsto t(0, \cos(\theta/2), \sin(\theta/2)). \quad (18.693)$$

L'image de  $u_1$  par cette réflexion est  $-u_1$ , c'est clair.

---

79. Les plus acharnées liront 4.6 pour comprendre la différence entre  $\{u_1, u_2, u_3\}$  et  $(u_1, u_2, u_3)$ .

Étudions en détail l'image de  $u_3$ . Nous devons démontrer que la droite  $D$  coupe le segment  $[u_3, A(u_3)]$  perpendiculairement en son milieu.

Nous considérons la base  $(u_2, u_3)$  du plan  $\text{Span}(u_2, u_3)$ . En utilisant les coordonnées dans cette base, nous avons  $u_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  et  $A(u_3) = \begin{pmatrix} \sin(\theta) \\ -\cos(\theta) \end{pmatrix}$ . Le milieu du segment  $[u_3, A(u_3)]$  est le point

$$M = \left( \frac{\sin(\theta)}{2}, \frac{1 - \cos(\theta)}{2} \right). \quad (18.694)$$

Les formules de duplication d'angle du corolaire 18.15 nous permettent d'écrire  $\sin(\theta)$  et  $\cos(\theta)$  en fonction de  $\sin(\theta/2)$  et  $\cos(\theta/2)$ , et donc d'exprimer le point  $M$  de la façon suivante :

$$M = \left( \cos(\theta/2) \sin(\theta/2), \frac{1 - (\cos^2(\theta/2) - \sin^2(\theta/2))}{2} \right) \quad (18.695a)$$

$$= (\cos(\theta/2) \sin(\theta/2), \sin^2(\theta/2)) \quad (18.695b)$$

$$= \sin(\theta/2) (\cos(\theta/2), \sin(\theta/2)). \quad (18.695c)$$

Ce point fait donc partie de la droite  $D$ . La droite  $D$  coupe le segment  $[u_3, A(u_3)]$  en son milieu.

En ce qui concerne l'orthogonalité, nous calculons le produit scalaire

$$(A(u_3) - u_3) \cdot \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = \sin(\theta) \cos(\theta/2) - (1 + \cos(\theta)) \sin(\theta/2) = 0 \quad (18.696)$$

où nous avons encore utilisé les duplications d'angles et le fait que  $1 = \cos^2(\theta/2) + \sin^2(\theta/2)$  (lemme 18.4).

(3) Si  $\epsilon = -1$  et  $\lambda = 1$  Alors la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{pmatrix}. \quad (18.697)$$

C'est la symétrie orthogonale par le plan engendré par  $u_1$  et  $v = \cos(\theta/2)u_2 + \sin(\theta/2)u_3$ .

Le vecteur  $u_1$  est bien évidemment préservé par  $A$ . En ce qui concerne le vecteur  $v$ ,

$$A(v) = \cos(\theta/2) \begin{pmatrix} 0 \\ \cos(\theta) \\ \sin(\theta) \end{pmatrix} + \sin(\theta/2) \begin{pmatrix} 0 \\ -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = \begin{pmatrix} 0 \\ \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = v. \quad (18.698)$$

Nous avons sauté quelques étapes de calcul mettant en scène les formules de duplication d'angle<sup>80</sup> : exprimer  $\cos(\theta) = \cos^2(\theta/2) - \sin^2(\theta/2)$  et  $\sin(\theta) = 2 \cos(\theta/2) \sin(\theta/2)$ .

Pour achever, nous devons trouver un vecteur  $w$  perpendiculaire au plan, et montrer qu'il est envoyé par  $A$  sur  $-w$ . Un vecteur  $w = xu_1 + yu_2 + zu_3$  est perpendiculaire au plan si les deux égalités suivantes sont satisfaites :

$$(\cos(\theta/2)u_2 + \sin(\theta/2)u_3) \cdot (xu_1 + yu_2 + zu_3) = 0 \quad (18.699a)$$

$$u_1 \cdot (xu_1 + yu_2 + zu_3) = 0. \quad (18.699b)$$

Nous avons immédiatement  $x = 0$  et ensuite la relation

$$y \cos(\theta/2) + z \sin(\theta/2) = 0. \quad (18.700)$$

Nous devons prouver que  $A(w) = -w = -yu_2 - zu_3$ , c'est-à-dire que

$$y \begin{pmatrix} 0 \\ \cos(\theta) \\ \sin(\theta) \end{pmatrix} + z \begin{pmatrix} 0 \\ \sin(\theta) \\ -\cos(\theta) \end{pmatrix} = \begin{pmatrix} 0 \\ -y \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -z \end{pmatrix}. \quad (18.701)$$

80. Corolaire 18.15.

La première ligne est immédiate. Nous faisons la seconde, et nous laissons au lecteur la troisième<sup>81</sup>. En utilisant les formules de duplication d'angle,

$$y \cos(\theta) + z \sin(\theta) = y(\cos^2(\theta/2) - \sin^2(\theta/2)) + 2z \cos(\theta/2) \sin(\theta/2) \quad (18.702a)$$

$$= y(2 \cos^2(\theta/2) - \sin^2(\theta/2) - \cos^2(\theta/2)) + 2z \cos(\theta/2) \sin(\theta/2) \quad (18.702b)$$

$$= -y + 2 \cos(\theta/2)(y \cos(\theta/2) + z \sin(\theta/2)) \quad (18.702c)$$

$$= -y \quad (18.702d)$$

où pour la dernière ligne, nous avons utilisé la relation (18.700).

(4) Si  $\epsilon = 1$  et  $\lambda = -1$  Alors la matrice est

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (18.703)$$

Cela est la composition entre la symétrie de plan  $\text{Span}(u_2, u_3)$  et la rotation d'angle  $\theta$  dans ce plan.

### 18.12.2 Sous-groupes finis de $SO(3)$

LEMooWIMMooXOCfSt

**Lemme 18.222** ([1]).

*Points fixes pour  $SO(3)$ .*

(1) *Tout élément de  $SO(3)$  possède une droite de points fixes.*

(2) *Tout élément non trivial de  $SO(3)$  possède une seule droite de points fixes.*

*Démonstration.* Le polynôme caractéristique d'un élément de  $SO(3)$  est de degré trois et possède donc (en comptant les multiplicités), trois racines dont une réelle par le corolaire 12.93. Comme nous sommes en dimension impaire, le coefficient du terme de degré 3 est  $-1$  et le polynôme caractéristique de  $g \in SO(3)$  s'écrit

$$\chi_g(X) = -(X - \lambda_1)(X - \bar{\lambda}_1)(X - s) \quad (18.704)$$

avec  $s = \pm 1$  que nous allons tout de suite fixer. Nous savons que  $\det(g) = \chi_g(0)$  mais aussi que  $\det(g) = 1$ . Donc

$$1 = \det(g) = \chi_g(0) = \lambda_1 \bar{\lambda}_1 s = s. \quad (18.705)$$

Tout cela pour dire que tout élément de  $SO(3)$  possède une valeur propre égale à 1, et donc une droite de points fixes.

Pour continuer, supposons que  $g$  possède deux droites distinctes de points fixes. En particulier  $g$  fixe un plan. Une base orthonormée de  $\mathbb{R}^3$  peut être choisie en prenant deux vecteurs  $e_1, e_2$  dans ce plan et un vecteur  $e_3$  perpendiculaire au plan.

Puisque  $g$  est une isométrie, la base reste orthonormée sous l'action de  $g$ . Donc  $g$  a pour matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}. \quad (18.706)$$

Pour que le déterminant soit 1, il faut que la matrice soit l'identité. □

PROPooBHPNooHPigwH

**Proposition 18.223** ([1, 620, 136, 621, 555]).

*Les sous-groupes finis de  $SO(3)$  sont :*

81. Je n'ai pas vérifié, alors faites attention.

- (1) les groupes cycliques  $\mathbb{Z}/n\mathbb{Z}$ ,  
 (2) les groupes diédraux  $D_n$ ,  
 (3) le groupe alterné  $A_4$ ,  
 (4) le groupe alterné  $A_5$ ,  
 (5) le groupe symétrique  $S_4$ .

*Démonstration.* Soit  $G$ , un sous-groupe fini de  $SO(3)$ . Par la proposition 8.45, les éléments de  $G$  sont des isométries de  $\mathbb{R}^3$ , et le lemme 18.222 dit que tout élément de  $G$  possède une droite de points fixes.

Un point de la sphère unité fixé par  $g \in G$  est un **pôle** de  $g$ . Nous nommons  $\Omega$  l'ensemble des pôles des éléments non triviaux de  $G$ .

- (1) **Une action** Le groupe  $G$  agit sur  $\Omega$ . En effet si  $x \in \Omega$ , alors  $x$  est fixé par un élément  $g$ . Montrons que  $h(x)$  est également fixé par un élément de  $G$ . Pas dur :  $(hgh^{-1})h(x) = h(x)$ ; donc  $h(x)$  est un pôle de  $hgh^{-1}$ .  
 (2) **Les fixateurs sont cycliques** Nous montrons à présent que pour tout  $x \in \Omega$ , le sous-groupe  $\text{Fix}(x)$  est cyclique. Soit donc  $x \in \Omega$ , le plan orthogonal  $\sigma = \text{Span}(x)^\perp$  et  $h \in \text{Fix}(x)$ . Nous avons  $h(\sigma) = \sigma$ . En effet si  $y \in \sigma$  nous avons

$$0 = y \cdot x = h(y) \cdot h(x) = h(y) \cdot x, \quad (18.707)$$

donc  $h(y)$  est perpendiculaire à  $x$ . L'inclusion inverse se démontre de même : si  $y \in \sigma$  alors  $y = h(h^{-1}(y))$  alors que  $h^{-1}(y) \in \sigma$ .

La restriction de  $h$  à  $\sigma$  est une isométrie de  $\sigma$ . Prenant une isométrie  $f: \sigma \rightarrow \mathbb{R}^2$ , l'application

$$\begin{aligned} \varphi: \text{Fix}(x) &\rightarrow SO(2) \\ h &\mapsto f \circ h \circ f^{-1}. \end{aligned} \quad (18.708)$$

est un morphisme injectif de groupes. En effet nous avons d'une part

$$\varphi(hh') = f \circ h \circ h' \circ f^{-1} = fhf^{-1}fh'f^{-1} = \varphi(h)\varphi(h'), \quad (18.709)$$

d'où le morphisme. Et d'autre part, si  $\varphi(h) = \varphi(h')$  alors  $f \circ h \circ f^{-1} = f \circ h' \circ f^{-1}$ , qui donne immédiatement  $h = h'$ .

Nous en déduisons que  $\text{Fix}(x)$  est isomorphe à un sous-groupe de  $SO(2)$  (l'image de  $\varphi$ ). Le lemme 18.141 en fait un groupe cyclique.

- (3) **Taille des fixateurs** Soient  $\Omega_i$  les orbites. Si  $x, y \in \Omega_i$  alors nous montrons que  $|\text{Fix}(x)| = |\text{Fix}(y)|$  avec la bijection

$$\begin{aligned} \varphi: \text{Fix}(x) &\rightarrow \text{Fix}(y) \\ h &\mapsto g^{-1}hg \end{aligned} \quad (18.710)$$

où  $g$  est choisie de façon à avoir  $y = g(x)$  (possible parce que  $x$  et  $y$  sont dans la même orbite). Cela est surjectif parce que si  $k \in \text{Fix}(x)$  alors  $k = \varphi(gkg^{-1})$  et l'on vérifie que  $gkg^{-1} \in \text{Fix}(y)$ . L'application  $\varphi$  est également injective parce que si  $ghg^{-1} = gh'g^{-1}$  alors  $h = h'$ .

- (4) **Un peu de notations** Puisque tous les fixateurs des éléments d'une orbite ont la même taille (finie), nous pouvons noter

$$n_i = |\text{Fix}(x_i)| \quad (18.711)$$

pour  $x_i \in \Omega_i$ . Nous notons également  $r$  le nombre d'orbites de  $G$ .

La formule de Burnside du théorème 2.41, avec les notations d'ici, donne

$$r = \frac{1}{|G|} \sum_{g \in G} |\text{Fix}(g)|. \quad (18.712)$$

- (5) **Une belle formule** Soit l'ensemble

$$A = \{(g, x) \text{ tel que } g \in G \setminus \{e\}, x \in \text{Fix}(g)\} \quad (18.713)$$

où par  $\text{Fix}(g)$  nous entendons les pôles de  $G$  fixés par  $g$ .

Il y a  $|G| - 1$  possibilités pour la composante  $g$ , mais chaque élément  $g \neq e$  possède exactement deux pôles, donc l'ensemble  $A$  contient exactement  $2(|G| - 1)$  éléments.

Nous pouvons calculer le nombre d'éléments dans  $A$  d'une autre façon : pour chaque  $x \in \Omega$  nous avons  $|\text{Fix}(x)| - 1$  éléments de  $G \setminus \{e\}$  qui fixent  $x$ . Donc

$$|A| = \sum_{x \in \Omega} (|\text{Fix}(x)| - 1). \quad (18.714)$$

Mais  $|\text{Fix}(x)|$  est constant sur les orbites. Nous coupons donc la somme sur  $\Omega$  en plusieurs sommes sur les orbites  $\Omega_i$  :

$$|A| = \sum_{i=1}^r \sum_{x \in \Omega_i} (|\text{Fix}(x)| - 1) = \sum_i |\Omega_i|(n_i - 1). \quad (18.715)$$

En égalisant les deux façons de calculer  $|A|$ , nous déduisons la formule

$$2(|G| - 1) = \sum_i |\Omega_i|(n_i - 1). \quad (18.716)$$

Nous utilisons ensuite la relation orbite-stabilisateur, proposition 2.34 :  $|\text{Fix}(x_i)||\Omega_i| = |G|$  ; la formule (18.716) devient

$$2(|G| - 1) = \sum_i |G| - \sum_i \frac{|G|}{n_i} = r|G| + |G| \sum_i \frac{1}{n_i}, \quad (18.717)$$

ou encore, en simplifiant par  $|G|$  :

$$2 - \frac{2}{|G|} = r - \sum_i \frac{1}{n_i} = \sum_{i=1}^r \left(1 - \frac{1}{n_i}\right). \quad (18.718)$$

Nous pouvons aussi repartir de (18.716) et sommer de façon plus simple  $\sum_i |\Omega_i| = |\Omega|$  pour obtenir

$$2(|G| - 1) = r|G| - |\Omega| \quad (18.719)$$

où  $\Omega$  est l'ensemble des pôles de  $G \setminus \{e\}$ .

- (6) **Quelles sont les possibilités ?** Les nombres  $|G|$ ,  $r$  et  $n_i$  sont des entiers. Nous allons voir qu'il n'y a pas des centaines de possibilités pour satisfaire la relation (18.718). D'abord, pour toute valeur de  $|G|$  (strictement plus grande que 1),

$$1 \leq 2 - \frac{2}{|G|} < 2. \quad (18.720)$$

Ensuite, si  $g$  fixe  $x$  alors  $g^{-1}$  fixe également  $x$ , de sorte que  $n_i = |\text{Fix}(x_i)| \geq 2$  pour tout  $i$ . Donc tous les termes dans la somme à droite de (18.718) sont dans  $[\frac{1}{2}, 1[$ . Nous avons donc au minimum deux termes, et au maximum trois. Autrement dit :  $r = 2$  ou  $r = 3$ .

- (7) **Si  $r = 2$**  Le plus simple est de repartir de (18.719). En posant  $r = 2$  nous trouvons tout de suite  $|\Omega| = 2$ . Il y a donc exactement deux pôles pour l'action de  $G$  sur la sphère unité.

Tous les éléments de  $G$  laissent donc le même axe invariant et  $G$  est un sous-groupe des isométries du plan qui lui est perpendiculaire. Autrement dit,  $G$  est un sous-groupe fini de  $SO(2)$  et donc cyclique par le lemme 18.141.

Nous étudions à présent le cas  $r = 3$ . Vu que ça va être un peu long, nous sautons un niveau d'indentation.



- (1) **Les possibilités pour  $r = 3$**  Nous devons voir les solutions entières  $(n_1, n_2, n_3, |G|)$  de

$$2 - \frac{2}{|G|} = 3 - \frac{1}{n_1} - \frac{1}{n_2} - \frac{1}{n_3} < 2. \quad (18.721)$$

Il faut en particulier que

$$\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} > 1, \quad (18.722)$$

ce qui signifie qu'au moins un des  $n_i$  doit être 1 ou 2, mais qu'il n'est pas possible que tous les  $n_i$  soient plus grands ou égaux à 3. Puisque  $n_i = |\text{Fix}(x_i)| \geq 2$ , nous en déduisons qu'au moins un des  $n_i$  doit valoir 2. Nous posons donc  $n_1 = 2$ .

De plus, nous savons que les  $n_i$  doivent diviser  $|G|$ . Donc  $|G|$  est pair.

- (2) **Si  $n_2 = 2$**  Nous sommes dans le cas  $r = 3$ ,  $n_1 = 2$ ,  $n_2 = 2$ . Nous avons

$$\frac{1}{n_3} = \frac{2}{|G|}, \quad (18.723)$$

mais aussi  $n_3 = |G|/|\Omega_3|$  d'où nous déduisons que  $|\Omega_3| = 2$ . Nous avons donc une orbite à deux éléments. Soit  $\Omega_3 = \{x, y\}$  avec  $x \neq y$ .

Le groupe  $\text{Fix}(x)$  est un groupe à  $|G|/2$  éléments. Il est donc normal par le lemme 3.24. Si  $g \in G$  est tel que  $g(x) = y$  alors nous avons  $\text{Fix}(y) = g \text{Fix}(x) g^{-1}$ , mais comme  $\text{Fix}(x)$  est normal nous avons  $\text{Fix}(x) = \text{Fix}(y)$ . Donc tous les éléments de  $\text{Fix}(x)$  fixent  $x$  et  $y$ . Le groupe  $\text{Fix}(x)$  est donc un sous-groupe de  $\text{SO}(2)$  et est cyclique, comme vu plus haut.

Mais de plus nous avons forcément  $y = -x$  parce qu'un élément de  $G$  qui fixe un point fixe également l'opposé. Vu que  $\Omega_3 = \{x, -x\}$ , il existe  $s \in G$  tel que  $s(x) = -x$ . Évidemment,  $s$  n'est pas dans  $\text{Fix}(x)$  et les points fixes de  $s$  ne sont pas parmi  $x$  et  $-x$ . Donc l'élément  $s^2$  a au moins 4 points fixes : les deux de  $s$  ainsi que  $x$  et  $-x$ . Il a donc au moins deux droites de points fixes, et est donc l'identité :  $s^2 = e$ .

De plus, vu que  $s(y)$  doit être égal soit à  $x$  soit à  $y$ , et que  $s(x) = y$ , l'injectivité de  $s$  donne  $s(y) = x$ .

Soit  $a$ , un générateur de  $\text{Fix}(x)$ . Nous allons montrer que  $G = \text{gr}(s, sa)$ . Nous avons déjà

$$(sa)(x) = s(x) = y \quad (18.724a)$$

$$(sa)(y) = s(y) = x. \quad (18.724b)$$

Donc  $sa$  inverse  $x$  et  $y$ . Mais  $sa$  a ses propres deux points fixes (qui ne sont ni  $x$  ni  $y$ ). L'élément  $(sa)^2$  a donc quatre points fixes sur la sphère unité :  $x$ ,  $y$  et les deux de  $sa$ . Nous en déduisons que  $(sa)^2 = e$ .

Nous nous souvenons que  $a$  est un générateur  $\text{Fix}(x)$ . Mais  $a = s \cdot sa$ , donc  $a^k = (ssa)^k$ . Nous en déduisons que  $\text{gr}(s, sa)$  contient au moins  $\text{Fix}(x)$ .

D'autre part si  $h$  et  $h'$  sont des éléments distincts dans  $\text{Fix}(x)$ , alors  $sh$  et  $sh'$  sont des éléments distincts de  $\text{gr}(s, sa)$  qui ne sont pas dans  $\text{Fix}(x)$ . Autrement dit, la partie

$$A = \{sh \text{ tel que } h \in \text{Fix}(x)\} \quad (18.725)$$

est une partie de même cardinal que  $\text{Fix}(x)$  tout en n'ayant aucune intersection avec  $\text{Fix}(x)$  (note : l'identité n'est pas dans  $A$ ). Mais  $|\text{Fix}(x)| = |G|/2$ , donc  $A \cup \text{Fix}(x) = G$ . Et justement  $A \cup G \subset \text{gr}(s, sa)$ . Nous en déduisons que  $\text{gr}(s, sa) = G$ .

Le théorème 18.185 nous assure que le groupe  $G$  est alors le groupe diédral parce que les éléments  $s$  et  $sa$  vérifient les relations données en 18.179.

- (3) **Si  $r = 3$ , les autres cas possibles** Nous repartons de (18.718) en posant  $r = 3$ . Nous obtenons ceci :

$$1 + \frac{2}{|G|} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}. \quad (18.726)$$

Nous avons déjà vu que  $n_1 = 2$  était obligatoire, et que tous les cas où deux des  $n_i$  sont égaux à 2 sont déjà couverts. Donc  $n_2$  et  $n_3$  valent 3 ou plus.

Nous trions les  $n_i$  dans l'ordre croissant. Si  $n_2 = 4$  ou plus, alors  $n_3$  vaut 4 ou plus. Mais

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 1 < 1 + \frac{2}{|G|}. \quad (18.727)$$

Donc  $n_3 = 3$  est obligatoire. Nous avons alors l'inégalité suivante qui restreint  $n_3$  :

$$\frac{1}{n_3} = \frac{1}{6} + \frac{3}{|G|} > \frac{1}{6}. \quad (18.728)$$

Donc  $n_3$  est 3, 4 ou 5.

Les derniers cas à couvrir sont :

- $(n_1, n_2, n_3) = (2, 3, 3)$ . Dans ce cas,  $\frac{7}{6} = 1 + \frac{2}{|G|}$ , donc  $|G| = 12$ .
- $(n_1, n_2, n_3) = (2, 3, 4)$ . Dans ce cas,  $|G| = 24$ .
- $(n_1, n_2, n_3) = (2, 3, 5)$ . Dans ce cas,  $|G| = 60$ .

(4) **Le cas  $(2, 3, 3)$**

Nous utilisons les relations  $n_i|\Omega_i| = |G|$  pour connaître la taille des orbites. Nous avons :

$$(4a) \quad 2|\Omega_1| = 12, \text{ donc } |\Omega_1| = 6,$$

$$(4b) \quad 3|\Omega_2| = 12, \text{ donc } |\Omega_2| = 4,$$

$$(4c) \quad 3|\Omega_3| = 12, \text{ donc } |\Omega_3| = 4.$$

Nous avons  $G \cdot \Omega_2 = \Omega_2$ . D'une part parce que, par définition d'une orbite,  $G \cdot \Omega_2 \subset \Omega_2$ , et d'autre part parce que si  $x \in \Omega_2$ , alors  $g^{-1}(x) \in \Omega_2$  et  $g(g^{-1}(x)) = x$ ; donc  $\Omega_2$  est bien dans l'image de  $\Omega_2$  par  $G$ . Nous avons donc un morphisme  $s: G \rightarrow S_{\Omega_2}$  que nous allons immédiatement voir comme

$$s: G \rightarrow S_4 \quad (18.729)$$

où  $S_4$  est le groupe des permutations de  $\{1, 2, 3, 4\}$ .

Voyons que  $s$  est injective. Si  $s(g) = s(h)$ , alors  $s(gh^{-1}) = \text{Id}$ . Autrement dit, l'élément  $sh^{-1}$  de  $G$  est l'identité sur  $\Omega_2$  qui contient 4 éléments. Fixant 4 points (au moins), l'élément  $sh^{-1}$  est l'identité. Par conséquent

$$s: G \rightarrow s(G) \subset S_4 \quad (18.730)$$

est un isomorphisme entre  $G$  et un sous-groupe de  $S_4$ . Mais  $|G| = 12$  et  $|S_4| = 24$ , donc  $G$  est d'indice deux dans  $S_4$  et est donc le groupe alterné  $A_4$  par la proposition 5.45(3).

(5) **Le cas  $(2, 3, 4)$**  Nous avons  $|G| = 24$  et les orbites ont pour taille :

$$— \quad 2|\Omega_1| = 24, \text{ donc } |\Omega_1| = 12,$$

$$— \quad 3|\Omega_2| = 24, \text{ donc } |\Omega_2| = 8,$$

$$— \quad 4|\Omega_3| = 24, \text{ donc } |\Omega_3| = 6.$$

(5a)  **$\Omega_2$  vient par paires** Soit  $x \in \Omega$  tel que  $|\text{Fix}(x)| = 3$ . Alors  $x \in \Omega_2$  parce que  $x$  est forcément dans un des  $\Omega_i$  et tout élément  $x_i$  de  $\Omega_i$  vérifie  $|\text{Fix}(x_i)| = n_i$ . Mais comme les éléments de  $\text{SO}(3)$  sont des applications linéaires, ceux qui fixent  $x$  fixent également  $-x$ . Cela pour dire que si  $x \in \Omega_2$ , alors  $-x \in \Omega_2$ . Nous avons donc quatre éléments distincts  $a_1, a_2, a_3$  et  $a_4$  tels que

$$\Omega_2 = \{\pm a_1, \pm a_2, \pm a_3, \pm a_4\}. \quad (18.731)$$

(5b) **Action sur les couples** Nous prétendons que  $G$  agit sur l'ensemble des couples  $\{\pm a_i\}$ . C'est encore la linéarité qui joue : l'élément  $g(a_i)$  est forcément un des  $\pm a_k$  (éventuellement  $k = i$ ). Si  $g(a_i) = a_k$ , alors  $g(-a_i) = -a_k$ . Autrement dit, pour tout  $i$ , il existe un  $k$  tel que  $g(\{a_i, -a_i\}) = \{a_k, -a_k\}$ . Cette association  $i \mapsto k$  est bijective (sinon  $g$  ne serait pas bijective), et fournit donc un morphisme de groupes

$$s: G \rightarrow S_4. \quad (18.732)$$

- (5c)  **$s$  est injective** Nous prouvons à présent que  $s(g) = \text{Id}$  si et seulement si  $g = e$ . Dans un sens c'est évident :  $g(e) = \text{Id}$ . Dans l'autre sens, nous devons prouver que si  $g(a_i) \in \pm a_i$  pour tout  $i$  alors  $g = e$ .

Si  $g(a_i) = a_i$  pour tout  $i$ , alors  $g$  stabilise 4 points et l'affaire est pliée. Nous supposons qu'au moins un des  $a_i$  n'est pas stabilisé par  $g$ . Pour fixer les idées nous disons que c'est  $a_1$ . Nous avons donc  $g(a_1) = -a_1$ . (oui :  $g(a_1) = -a_1$  et non  $\pm a_k$  pour un autre  $k$  parce que nous sommes sous l'hypothèse que  $g$  stabilise les couples)

L'élément  $g^2$  fixe tout  $\Omega_2$ ; donc  $g^2 = e$ . Nommons  $\pm b$  les points fixes de  $g$ . Si  $b \in \Omega_2$  alors  $|\text{Fix}(b)| = 3$ , c'est-à-dire que les éléments de  $G$  qui fixent  $b$  sont dans un groupe d'ordre 3, et le corolaire 2.15 nous indique que ces éléments ne peuvent être que d'ordre 1 ou 3, pas deux. Nous en déduisons que  $b$  n'est pas dans  $\Omega_2$  et donc que  $g(a_i) = -a_i$  pour tout  $i$ .

Jusqu'à présent nous avons prouvé que si  $g \in \ker(s)$  est non trivial, alors  $g(a_i) = -a_i$  pour tout  $i$ .

Soit maintenant  $h \in G$ . Comme  $\Omega_2$  est une orbite,  $h(a_i) \in \Omega_2$  et nous notons  $h(a_i) = \epsilon a_k$  avec  $\epsilon = \pm 1$  et éventuellement  $k = i$  ou éventuellement pas. Nous avons :

$$(h^{-1}gh)(a_i) = \epsilon(h^{-1}g)(a_k) = -\epsilon h^{-1}(a_k) = -\epsilon^2 a_i = -a_i. \quad (18.733)$$

Donc  $g$  et  $h^{-1}gh$  ont même restriction à  $\Omega_2$ . En particulier  $h^{-1}ghg^{-1}$  est l'identité sur  $\Omega_2$  et est donc l'identité.

Pour tout  $h$  nous avons  $g = h^{-1}gh$ . Les points fixes de  $h^{-1}gh$  sont  $\pm h^{-1}(b)$ , mais aussi  $\pm b$ . Nous avons donc égalité d'ensemble  $\{h(b), -h(b)\} = \{b, -b\}$  pour tout  $h \in G$  (notez le changement de notation  $h \rightarrow h^{-1}$ ). Cela signifie que  $\{b, -b\}$  est une orbite de  $G$ . Mais on a pas d'orbites de cardinal deux; contradiction. Nous en déduisons que  $e$  est l'unique élément de  $\ker(s)$ .

- (5d) **Conclusion** La partie  $s(G)$  est un sous-groupe de  $S_4$  isomorphe à  $G$ . Mais au niveau des cardinaux,  $|G| = 24$  en même temps que  $|S_4| = 24$ . Donc  $G \simeq s(G) \simeq S_4$ .

Nous passons au cas  $(2, 3, 5)$ , et comme ça va être long et douloureux<sup>82</sup>, nous sautons un niveau d'indentation.

Au niveau du cardinal de  $G$ ,

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} = 1 + \frac{2}{|G|}, \quad (18.734)$$

donc  $|G| = 60$ . Et pour les orbites,  $|\Omega_1| = 30$ ,  $|\Omega_2| = 20$ ,  $|\Omega_3| = 12$ .

La proposition 5.55 nous indique que le seul groupe simple d'ordre 60 est le groupe  $A_5$ . Nous allons donc nous atteler à prouver que  $G$  est simple. Vous êtes prêts ?

- (1) **Fixateurs et ordres** Tous les éléments de  $G$  sont dans un fixateur de type  $\text{Fix}(x)$ , et comme l'ordre d'un élément divise l'ordre du groupe (corolaire 2.15), tous les éléments de  $G$  ont un ordre 2, 3 ou 5. Nous sommes dans un cas très particulier parce que

- Les trois nombres 2, 3 et 5 sont des nombres premiers distincts. Donc « diviser  $n_i$  » signifie pratiquement « être égal à  $n_i$  », surtout lorsqu'on parle de l'ordre d'un élément, qui ne peut pas être 1.
- Il existe une seule orbite de chaque taille.

Nous notons  $G(n_i)$  l'ensemble des éléments de  $G$  d'ordre  $n_i$ . Les parties  $G(n_i)$  ne contiennent pas l'identité.

- (2)  **$g \in G(n_i)$  implique  $\text{Fix}(g) \subset \Omega_i$**  Si  $g \in G(n_i)$  et  $x \in \text{Fix}(g)$  alors  $x \in \Omega_i$ . En effet  $x \in \text{Fix}(g)$  signifie  $g(x) = x$  et donc aussi  $g \in \text{Fix}(x)$ . Donc l'ordre de  $g$  divise  $|\text{Fix}(x)|$ , alors que l'ordre de  $g$  est  $n_i$  et que les possibilités pour  $|\text{Fix}(x)|$  sont exactement les  $n_i$ , lesquels sont premiers entre eux. Donc  $|\text{Fix}(x)| = n_i$  et  $x \in \Omega_i$ .

82. Mais pas autant que le théorème 32.36, cependant.

- (3)  $|\text{Fix}(x)| = n_i$  **implique**  $x \in \Omega_i$  Comme plus haut,  $g \in \text{Fix}(x)$  implique que l'ordre de  $g$  divise  $n_i$  et est donc égal à  $n_i$ . Autrement dit,  $g \in G(n_i)$ . De plus  $g \in \text{Fix}(x)$  implique  $x \in \text{Fix}(g)$ . Par le cas juste au-dessus nous déduisons  $x \in \Omega_i$ .
- (4)  **$a$  et  $-a$  dans la même orbite**  
 Nous avons évidemment  $\text{Fix}(a) = \text{Fix}(-a)$  parce que les éléments de  $G$  sont des applications linéaires. Si  $|\text{Fix}(a)| = n_i$  alors  $a \in \Omega_i$  et aussi  $|\text{Fix}(-a)| = |\text{Fix}(a)| = n_i$  et encore  $-a \in \Omega_i$ .
- (5) **Nombre de  $\text{Fix}(x_i)$**  Soient  $a, b \in \Omega_i$ . Nous avons  $|\text{Fix}(a)| = |\text{Fix}(b)| = n_i$  et  $\text{Fix}(a) = \text{Fix}(b)$  si et seulement si  $b = -a$  parce qu'un élément qui fixe  $a$  et  $b$  fixe automatiquement  $a, b, -a$ , et  $-b$ . Aucun élément non trivial ne peut fixer quatre points distincts. Autrement dit,

$$\text{Fix}(a) \cap \text{Fix}(b) = \begin{cases} \text{Fix}(a) & \text{si } a = \pm b \\ \{e\} & \text{sinon.} \end{cases} \quad (18.735)$$

Chaque élément  $x_i \in \Omega_i$  a son fixateur (il y en aurait  $|\Omega_i| = 60/n_i$ ), mais ces fixateurs sont égaux deux à deux, donc il y a seulement  $\frac{60}{2n_i}$  groupes distincts de la forme  $|\text{Fix}(x_i)|$  avec  $x_i \in \Omega_i$ .

- (6) **Récapitulatif** En reprenant ce que nous venons de dire avec  $i = 1, 2, 3$  nous trouvons :

(6a)  $n_1 = 2$ , avec  $|\Omega_1| = 30$  et 15 groupes du type  $\text{Fix}(x_1)$  avec  $x_1$  parcourant  $\Omega_1$ .

(6b)  $n_2 = 3$ , avec  $|\Omega_2| = 20$  et 10 groupes du type  $\text{Fix}(x_2)$  avec  $x_2$  parcourant  $\Omega_2$ .

(6c)  $n_3 = 5$ , avec  $|\Omega_3| = 12$  et 6 groupes du type  $\text{Fix}(x_3)$  avec  $x_3$  parcourant  $\Omega_3$ .

Un élément non trivial de  $G$  se trouve forcément dans un et un seul de ces sous-groupes. Plus précisément, si  $g \in G(n_i)$  alors  $g$  est dans un des  $\text{Fix}(x_i)$  avec  $x_i \in \Omega_i$ .

Comptons pour être sûr de ne pas s'être trompé. Chacune des lignes décrit 30 éléments de  $G$ ; par exemple, la seconde ligne donne 10 groupes de taille  $|\text{Fix}(x_2)| = n_2 = 3$ . Mais tous ces groupes ont pour intersection exactement  $\{e\}$ . Donc le comptage des éléments se fait comme suit :

$$3 \times 30 - 15 - 10 - 6 + 1. \quad (18.736)$$

Le dernier  $+1$  est parce que nous aurions décompté l'identité une fois de trop. Bref, on a bien 60 éléments comme il se doit.

- (7) **Un ensemble à calculer deux fois** Soient les ensembles  $A_2, A_3$  et  $A_5$  définis par

$$A_i = \{(g, a) \in G(n_i) \times \Omega_i \text{ tel que } g(a) = a\} \quad (18.737)$$

où  $G(n_i)$  est la partie de  $G$  des éléments d'ordre  $n_i$ .

Nous avons

$$|A_i| = \sum_{g \in G(n_i)} |\text{Fix}(g) \cap \Omega_i|. \quad (18.738)$$

Mais les éléments de  $G(n_i)$  sont d'ordre  $n_i$ , et par ce que nous avons dit plus haut, tous les éléments de  $\text{Fix}(g)$  sont dans  $\Omega_i$ . Donc  $\text{Fix}(g) \cap \Omega_i = \text{Fix}(g)$ . Nous avons alors

$$|A_i| = \sum_{g \in G(n_i)} |\text{Fix}(g)| = 2|G(n_i)| \quad (18.739)$$

parce que  $|\text{Fix}(g)| = 2$  pour tout  $g$ .

En comptant  $|A_i|$  dans l'autre sens, nous avons

$$|A_i| = \sum_{x \in \Omega_i} |\text{Fix}(x) \cap G(n_i)| \quad (18.740)$$

Vu que  $x \in \Omega_i$ , les éléments de  $\text{Fix}(x)$  sont d'ordre  $n_i$ <sup>83</sup> (sauf  $e$ ), et comme  $G(n_i)$  est justement l'ensemble des éléments d'ordre  $n_i$  dans  $G$  nous avons  $\text{Fix}(x) \cap G(n_i) = \text{Fix}(x) \setminus \{e\}$ . Cela

83. Encore et toujours parce que les éléments de  $\text{Fix}(x)$  ont un ordre qui divise  $|\text{Fix}(x)| = n_i$  et que  $n_i$  est premier, et que nous avons exclu l'identité.

pour dire que

$$|A_i| = \sum_{x \in \Omega_i} (|\text{Fix}(x)| - 1) \quad (18.741a)$$

$$= \sum_{x \in \Omega_i} |\text{Fix}(x)| - \sum_{x \in \Omega_i} 1 \quad (18.741b)$$

$$= \sum_{x \in \Omega_i} n_i - |\Omega_i| \quad |\text{Fix}(x)| = n_i \text{ pcq } x \in \Omega_i \quad (18.741c)$$

$$= |\Omega_i|n_i - |\Omega_i| = |G| - |\Omega_i|. \quad (18.741d)$$

En égalisant cela à la valeur  $2|G(n_i)|$  déjà trouvée, nous déduisons les valeurs des  $|G(n_i)|$  :

$$|G(n_i)| = \frac{|G| - |\Omega_i|}{2}. \quad (18.742)$$

Nous avons alors

$$(7a) \quad |G(2)| = 15$$

$$(7b) \quad |G(3)| = 20$$

$$(7c) \quad |G(5)| = 24$$

- (8) **Les Sylow de  $G$**  Les  $p$ -Sylow sont définis en 5.6, et le super théorème qui répond à toutes les questions est le théorème 5.11. Dans notre cas, les diviseurs premiers de  $|G| = 60$  sont 2, 3 et 5. Il faut faire attention au 2 parce que sa plus haute puissance dans la décomposition de 60 est 4 et non 2. Nous avons :

(8a) Un 2-Sylow est un sous-groupe d'ordre 4.

(8b) Un 3-Sylow est un sous-groupe d'ordre 3.

(8c) Un 5-Sylow est un sous-groupe d'ordre 5.

Entre autres :

(8a) Les 10 sous-groupes  $\text{Fix}(x_2)$  avec  $x_2 \in \Omega_2$  sont des 3-Sylow de  $G$ .

(8b) Les 6 sous-groupes  $\text{Fix}(x_3)$  avec  $x_3 \in \Omega_3$  sont des 5-Sylow de  $G$ .

(8c) Les 15 sous-groupes  $\text{Fix}(x_1)$  avec  $x_1 \in \Omega_1$  sont d'ordre 2 et ne sont donc pas des 2-Sylow de  $G$ .

- (9) **Tous les 3-Sylow et les 5-Sylow** Nous avons déjà trouvé 10 3-Sylow et 6 5-Sylow. Nous montrons à présent qu'il n'y en a pas d'autres. Le théorème de Sylow 5.11(4) nous indique que le nombre  $n_3$  de 3-Sylow est :

- diviseur de 60,
- dans  $[1]_3$
- au moins 10.

Les diviseurs de 60 sont :

$$1, 5, 3, 15, 2, 10, 6, 30, 4, 20, 12, 60. \quad (18.743)$$

Le seul qui vérifie toutes les conditions est 10. Donc  $G$  possède seulement 10 3-Sylow et ils sont tous de la forme  $\text{Fix}(x_2)$  avec  $x_2 \in \Omega_2$ .

Même raisonnement pour les 5-Sylow : il faut

- diviseur de 60,
- dans  $[1]_5$
- au moins 6.

La seule possibilité est 6.

- (10) **Sous-groupe normal** Soit  $H$ , un sous-groupe normal de  $G$ . Notre but étant de prouver que  $G$  est simple, nous voulons prouver que  $H$  est soit  $\{e\}$  soit  $G$ . Nous supposons que  $H$  est non trivial, et nous allons prouver que  $H = G$ .

Le théorème de Lagrange 2.14(1) nous dit que  $|H|$  divise  $|G|$ . Le nombre  $|H|$  ne peut donc avoir que 2, 3 et 5 comme facteurs premiers. Avec une mention spéciale pour le 2 :  $|H|$  pourrait être divisible aussi par 4.

- (11) **Diviseurs de  $|H|$**

Soit un sous-groupe normal  $H$  de  $G$ . Puisque c'est un sous-groupe, son ordre divise celui de  $G$  (encore et toujours le théorème de Lagrange 2.14), et donc les facteurs premiers de  $|H|$  ne peuvent être que 2, 3 et 5.

- (12) **Si  $|H|$  est divisible en 3** Alors  $H$  contient au moins un 3-Sylow. Mais nous avons vu que les 3-Sylow de  $H$  sont les 3-Sylow de  $G$ . Donc  $H$  contient tous les 3-Sylow de  $G$ , parce que les 3-Sylow sont conjugués et  $H$  est normal.

Soit  $E$  l'ensemble des sous-groupes de  $H$ . Puisqu'il est normal,  $H$  agit sur  $E$  par conjugaison, et les 3-Sylow forment une orbite. Si  $\alpha$  est un 3-Sylow, la formule des classes (proposition 2.34(2)) nous donne

$$|H| = |\text{Fix}(\alpha)| |\mathcal{O}_\alpha|. \quad (18.744)$$

Mais l'orbite  $\mathcal{O}_\alpha$  de  $\alpha$  est l'ensemble des 3-Sylow, de sorte que  $|\mathcal{O}_\alpha| = 10$ . Donc  $|H|$  est divisible en 10.

Mais il y a pire :  $H$  contient au moins les 10 sous-groupes  $\text{Fix}(x_2)$  pour  $x_2 \in \Omega_2$ . Ce sont 10 groupes de  $|\text{Fix}(x_2)| = 3$  éléments. En décomptant  $e$  qui est dans l'intersection, cela fait

$$10 \times |\text{Fix}(x_2)| - 10 + 1 = 21 \quad (18.745)$$

éléments. Donc  $H$  contient au moins 21 éléments. Le nombre  $|H|$  est donc :

- diviseur de 60
- multiple de 10
- au moins 21.

Donc c'est 30 ou 60.

- (13) **Si  $|H|$  est divisible en 5** Le même raisonnement tient et  $|H|$  est 30 ou 60.

Nous restons avec les possibilités  $|H|$  égal à 2, 4, 30 ou 60.

- (1) **Si  $|H| = 4$**

Alors  $H$  contient au moins un 2-Sylow. Un 2-Sylow de  $H$  est un sous-groupe contenant 4 éléments qui sont d'ordre  $2^m$ . Le seul  $m$  possible dans  $G$  est  $m = 1$ . Vu qu'un 2-Sylow de  $H$  contient 4 éléments, nous sommes dans le cas où  $H$  est un 2-Sylow. Il est donc le seul 2-Sylow de  $H$  parce que  $H$  est normal, et que tous les 2-Sylow sont conjugués.

Mais tous les sous-groupes d'ordre 2 sont contenus dans un 2-Sylow. En particulier tous les 15 groupes  $\text{Fix}(x_1)$  sont dans l'unique 2-Sylow  $H$  qui est soi-disant d'ordre 4. IL y a là une belle impossibilité.

Donc, le cas  $|H| = 4$  est hors-concours.

- (2) **Si  $|H| = 2$**  Alors  $H = \{e, g\}$  avec  $g^2 = e$ . Si  $h \in G$ , l'élément  $hgh^{-1}$  ne peut être que  $e$  ou  $g$  (parce que  $H$  est normal). Le premier cas est  $g = e$ , et le second donne  $gh = hg$ . Donc  $g$  est dans le centre de  $G$  : il commute avec tous les éléments de  $G$ .

Comme  $g \in G(2)$ , nous savons que les éléments  $a \in \text{Fix}(g)$  sont forcément dans  $\Omega_1$  parce que les points dont les fixateurs sont formés d'éléments d'ordre 2 sont dans  $\Omega_1$ . Soit  $h \in G$ . Nous avons  $g = hgh^{-1}$  et donc aussi

$$(hgh^{-1})(h(a)) = hg(a) = h(a), \quad (18.746)$$

donc  $h(a)$  et  $-h(a)$  sont des points fixes de  $hgh^{-1}$ . Ce sont donc également des points fixes de  $g$ . Nous en déduisons que  $g$  a pour points fixes les points  $a$ ,  $-a$ ,  $h(a)$  et  $-h(a)$ . Puisque

$g$  n'est pas  $e$ , ces quatre points ne peuvent pas être distincts. Comme  $h(a)$  ne peut pas être  $-h(a)$ , nous avons forcément  $h(a) = \pm a$ .

Donc l'orbite de  $a$  ne contiendrait que 2 éléments. Pas possible.

- (3) Si  $|H| = 30$  À part  $|H| = 60$ , le dernier cas à traiter est  $|H| = 30$ . Nous rappelons obligeamment que

$$(3a) \quad |G(2)| = 15$$

$$(3b) \quad |G(3)| = 20$$

$$(3c) \quad |G(5)| = 24.$$

Si  $H$  possède 30 éléments, le théorème de Sylow dit que  $H$  contient au moins un 3-Sylow et un 5-Sylow, et donc tous. Puisque pour 3 et 5, les Sylow de  $H$  et de  $G$  sont les mêmes et bien identifiés, nous allons nous baser dessus. Le sous-groupe  $H$  contient tous les 3 et 5-Sylow, donc le comptage des éléments est :

$$10 \times |\text{Fix}(x_2)| + 6 \times |\text{Fix}(x_3)| - 15 = 45. \quad (18.747)$$

Nous aurions aussi pu ajouter  $+4 - 1$  pour compter au moins un 2-Sylow.

Donc dès que  $H$  compte 30 éléments, il en compte au moins 45. Autrement dit le cas  $|H| = 30$  est impossible.

□

## ⚠ Avertissement/question à la lectrice !! 18.224

La démonstration des groupes finis de  $\text{SO}(3)$  est longue. Je me demande si il n'y a pas moyen de faire plus court. Par exemple [620] utilise le théorème de Cauchy 3.22 que je n'utilise pas. D'autre part, toutes les références me semblent utiliser plus ou moins implicitement le fait que si le sous-groupe normal  $H$  contient un élément de  $G(n_i)$ , alors il les contiennent tous. J'avoue ne pas trop comprendre pourquoi.

## 18.13 Systèmes de coordonnées

SECooWTPRooZbOSzO

La trigonométrie nous offre de nouveaux systèmes de coordonnées qui peuvent se révéler pratiques dans certains cas : les coordonnées polaires sur  $\mathbb{R}^2$ , ainsi que les coordonnées cylindriques et sphériques sur  $\mathbb{R}^3$ .

### 18.13.1 Coordonnées polaires

#### 18.13.1.1 Ce que ça signifie intuitivement

On a vu qu'un point  $M$  dans  $\mathbb{R}^2$  peut être représenté par ses abscisses  $x$  et ses ordonnées  $y$ . Nous pouvons également déterminer le même point  $M$  en donnant un angle et une distance comme illustré sur la figure 18.10.

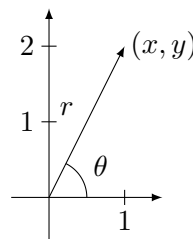


FIGURE 18.10: Un point en coordonnées polaires est donné par sa distance à l'origine et par l'angle qu'il fait avec l'horizontale.

LabelFigJWINooSfKCeA

Le même point  $M$  peut être décrit indifféremment avec les coordonnées  $(x, y)$  ou bien avec  $(r, \theta)$ .



**Remarque 18.225.**

L'angle  $\theta$  d'un point n'étant à priori défini qu'à un multiple de  $2\pi$  près, nous convenons de toujours choisir un angle  $0 \leq \theta < 2\pi$ . Par ailleurs l'angle  $\theta$  n'est pas défini si  $(x, y) = (0, 0)$ .

La coordonnée  $r$  est toujours positive.

Nous avons dans l'idée de définir  $r$  et  $\theta$  par les formules <sup>EqrthetaxyPoal</sup>

$$\begin{cases} x = r \cos(\theta) \\ y = r \sin(\theta). \end{cases} \quad \begin{matrix} (18.748a) \\ (18.748b) \end{matrix}$$

**18.13.1.2 Coordonnées polaires : le théorème**

THOooBETSoosXSQhdX

**Théorème 18.226** (Coordonnées polaires<sup>[1]</sup>).

Soit l'application

$$\begin{aligned} T: [0, \infty[ \times [0, 2\pi[ &\rightarrow \mathbb{R}^2 \\ (r, \theta) &\mapsto \begin{pmatrix} r \cos(\theta) \\ r \sin(\theta) \end{pmatrix}. \end{aligned} \quad (18.749)$$

(1) L'application  $T$  est surjective.

ITEMooNGOKooFCXmwy

(2) L'application

ITEMooMCIOooJiBvug

$$T: ]0, \infty[ \times [0, 2\pi[ \rightarrow \mathbb{R}^2 \setminus \{(0, 0)\} \quad (18.750)$$

est bijective.

ITEMooZFRGooQPDUtX

(3) En considérant la demi-droite  $D = \{(x, 0)\}_{x \geq 0}$ , l'application

$$T: ]0, \infty[ \times ]0, 2\pi[ \rightarrow \mathbb{R}^2 \setminus D \quad (18.751)$$

est un  $C^\infty$ -difféomorphisme<sup>84</sup>.

*Démonstration.* Une bonne partie de ce théorème est une conséquence de 18.59. La vraie nouveauté de ce théorème sera la régularité. Nous démontrons point par point.

(1) Pour (1). Soit  $a = (x, y) \in \mathbb{R}^2$ . Nous avons  $a/\|a\| \in S^1$ . Par la proposition 18.59, il existe  $\theta \in [0, 2\pi]$  tel que

$$\frac{a}{\|a\|} = (\cos(\theta), \sin(\theta)). \quad (18.752)$$

Alors  $a = \|a\|(\cos(\theta), \sin(\theta)) = T(\|a\|, \theta)$ . Voilà. L'application  $T$  est surjective.

(2) Pour (2). En ce qui concerne la surjectivité,

$$T(0, [0, 2\pi[) = \{(0, 0)\}. \quad (18.753)$$

Donc le point (1) donne le surjectif lorsque nous enlevons d'un côté les points avec  $r = 0$  et de l'autre le point  $(0, 0)$ .

Pour l'injectivité, nous supposons  $T(r_1, \theta_1) = T(r_2, \theta_2)$ . Puisque  $\|T(r, \theta)\| = r$ , nous avons tout de suite  $r_1 = r_2$ . Nous restons donc avec l'égalité

$$\begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{pmatrix} = \begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix}. \quad (18.754)$$

La proposition 18.59 dit alors que  $\theta_1 = \theta_2$ .

84. L'application est de classe  $C^\infty$  et son inverse est également de classe  $C^\infty$ . Le plus souvent, vous voulez seulement utiliser ce théorème dans le but de faire un changement de variables dans une intégrale ; vous n'avez donc besoin que d'un  $C^1$ -difféomorphisme.



- (3) Pour (3). L'application  $T$  est injective en tant que restriction d'une application injective. Pour le surjectif, soit  $a \in \mathbb{R}^2 \setminus D$ . Comme  $a \notin D$ , nous avons  $\|a\| \neq 0$  et il est légitime de dire, comme plus haut, qu'il existe  $\theta \in [0, 2\pi[$  tel que

$$\frac{a}{\|a\|} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}. \quad (18.755)$$

Ce  $\theta$  n'est pas zéro parce que  $\theta = 0$  donne le point  $(1, 0)$  qui est sur  $D$ .

En ce qui concerne l'inverse, nous n'allons pas nous lancer dans une étude subtile de la fonction (18.180); nous avons déjà démontré la continuité dans le lemme 18.66, et monter dans les dérivées nous semble un peu compliqué. Au lieu de cela, nous allons procéder en deux étapes :

- Prouver que  $T$  est de classe  $C^p$  pour tout  $p$  en invoquant seulement des théorèmes à proposition de différentielle,
- En déduire que  $T^{-1}$  est également  $C^p$  pour tout  $p$  en invoquant le théorème d'inversion locale 17.47.

Les applications  $(r, \theta) \mapsto r$ ,  $(r, \theta) \mapsto \sin(\theta)$  et  $(r, \theta) \mapsto \cos(\theta)$  sont de classe  $C^\infty$  grâce au lemme 12.273. Le lemme 12.294 sur la différentiabilité du produit montre alors que les fonctions  $T_1$  et  $T_2$  données par

$$T_1(r, \theta) = r \cos(\theta) \quad (18.756a)$$

$$T_2(r, \theta) = r \sin(\theta) \quad (18.756b)$$

sont différentiables<sup>85</sup>. Mieux, la proposition 12.297 montre que ces fonctions  $T_1$  et  $T_2$  sont de classe  $C^p$  pour tout  $p$ , c'est-à-dire qu'elles sont de classe  $C^\infty$ . Cela montre que les coordonnées polaires sont de classe  $C^\infty$ , et il faut encore parler de l'inverse.

En ce qui concerne la différentielle,

$$dT_{(r, \theta)}(u, v) = \begin{pmatrix} u \cos(\theta) - rv \sin(\theta) \\ u \sin(\theta) + rv \cos(\theta) \end{pmatrix}. \quad (18.757)$$

Donc la matrice de la différentielle est

$$dT_{(r, \theta)} = \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix}, \quad (18.758)$$

dont le déterminant est  $r$  (lemme 18.4 utilisé). Donc la différentielle en  $(r, \theta)$  est une application linéaire inversible parce que  $r \neq 0$  aux points que nous considérons. L'application  $dT_{(r, \theta)}$  est bicontinue parce que nous sommes en dimension finie. Tout cela pour dire que le théorème d'inversion local 17.47 s'applique, et  $T^{-1}$  est  $C^p$  dès que  $T$  est  $C^p$ .

Puisque  $T$  est de classe  $C^p$  pour tout  $p$ , l'inverse  $T^{-1}$  est également  $C^p$  pour tout  $p$ , c'est-à-dire que  $T^{-1}$  est de classe  $C^\infty$ .

□

### Définition 18.227.

Ce que nous appelons **les coordonnées polaires** est l'application

$$T: [0, \infty[ \times [0, 2\pi[ \rightarrow \mathbb{R}^2$$

$$(r, \theta) \mapsto \begin{pmatrix} r \cos(\theta) \\ r \sin(\theta) \end{pmatrix}. \quad (18.759)$$

du théorème 18.226(3). Selon les circonstances, nous considérons l'une ou l'autre des restrictions pour avoir une bijection ou un difféomorphisme.

<sup>85</sup>. Si vous voulez seulement avoir un  $C^1$ -difféomorphisme, calculez explicitement la différentielle et montrez que c'est continu. Vous n'avez pas à utiliser la proposition 12.297 ni rien des produits tensoriels.

**Exemple 18.228.**

Soit à calculer

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^2 + y^2}{x - y}. \quad (18.760)$$

Nous introduisons la fonction

$$\begin{aligned} f: \mathbb{R}^2 \setminus \{x = y\} &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \frac{x^2 + y^2}{x - y}. \end{aligned} \quad (18.761)$$

Une idée souvent fructueuse pour traiter ce genre de limite est de passer aux coordonnées polaires. Attention, si on veut faire les choses très explicitement, c'est un peu lourd en notations. Il s'agit de poser

$$\begin{aligned} g: ]0, \infty[ \times [0, 2\pi[ \setminus \{\mathbb{R} \times \{\frac{\pi}{4}\} \cup \mathbb{R} \times \{\frac{5\pi}{4}\}\} &\rightarrow \mathbb{R} \\ (r, \theta) &\mapsto \frac{r^2}{r(\cos(\theta) - \sin(\theta))}. \end{aligned} \quad (18.762)$$

Bon. À strictement parler, nous aurions pu dire que  $g$  est définie pour  $r = 0$ , mais vu que nous voulons seulement calculer la limite pour  $r \rightarrow 0$ , on n'a pas besoin de la valeur en zéro. De plus les coordonnées polaires ne sont pas bijectives en l'origine. Donc bon... on s'en passe.

Quel est le lien entre  $f$  et  $g$ ? Du point de vue du calcul, le lien est qu'on a remplacé  $x$  par  $r \cos(\theta)$  et  $y$  par  $r \sin(\theta)$ . Le vrai lien est l'égalité

$$g = f \circ T \quad (18.763)$$

où  $T$  est l'application de coordonnées polaires dont les principales propriétés sont données dans le théorème 18.226(2).

Soit un voisinage  $B((0,0), R)$  de  $(0,0)$  dans  $\mathbb{R}^2$ . Le but est de montrer que les valeurs  $f(B)$  se regroupent autour d'une valeur  $\ell$  lorsque  $R \rightarrow 0$ . Soyons plus précis et nommons  $\ell$  le candidat limite. Soit  $\epsilon > 0$ ; nous devons trouver  $R > 0$  tel que  $f(B((0,0), R)) \subset B(\ell, \epsilon)$ .

Pour  $R > 0$ , nous avons

$$B((0,0), R) = T([0, R[ \times [0, 2\pi[), \quad (18.764)$$

donc

$$f(B) = g([0, R[ \times [0, 2\pi[). \quad (18.765)$$

Soit  $r < R$ . Nous avons

$$\lim_{\theta \rightarrow \pi/4} g(r, \theta) = \infty. \quad (18.766)$$

Donc  $f(B)$  contient des valeurs arbitrairement grandes, quelle que soit la valeur de  $R$ . Il n'y a donc pas de limite possible.

Si vous voulez un argument un peu plus imagé, en voici un <sup>86</sup> basé sur une combinaison entre la méthode des coordonnées polaires et la méthode des chemins.

Certes pour chaque  $\theta$  nous avons  $\lim_{r \rightarrow 0} g(r, \theta) = 0$ , mais il ne faut pas en déduire trop vite que la limite  $\lim_{(x,y) \rightarrow (0,0)} g(x, y)$  vaut zéro parce que prendre la limite  $r \rightarrow 0$  avec  $\theta$  fixé revient à prendre la limite le long de la droite d'angle  $\theta$ .

Il n'est pas possible de majorer  $g(r, \theta)$  par une fonction ne dépendant pas de  $\theta$  parce que cette fonction tend vers l'infini lorsque  $\theta \rightarrow \pi/4$ . Est-ce que cela veut dire que la limite n'existe pas? Cela veut en tout cas dire que la méthode des coordonnées polaires ne parvient pas à résoudre l'exercice. Pour conclure, il faudra encore un peu travailler.

<sup>86</sup>. Qui satisfera tous vos professeurs, pourvu que vous ayez compris que ce qui se cache est une histoire de valeurs de  $f$  prises sur un voisinage de  $(0,0)$ .

Nous pouvons essayer de calculer le long d'un chemin plus général  $(r(t), \theta(t))$ . Choisissons  $r(t) = t$  puis cherchons  $\theta(t)$  de telle sorte à avoir

$$\cos \theta(t) - \sin \theta(t) = t^2. \quad (18.767)$$

Le mieux serait de résoudre cette équation pour trouver  $\theta(t)$ . Mais en réalité il n'est pas nécessaire de résoudre : montrer qu'il existe une solution suffit. Nous pouvons supposer que  $t^2 < 1$ . Pour  $\theta = \pi/4$  nous avons  $\cos(\theta) - \sin(\theta) = 0$  et pour  $\theta = 0$  nous avons  $\cos(\theta) - \sin(\theta) = 1$ . Le théorème des valeurs intermédiaires nous enseigne alors qu'il existe une valeur de  $\theta$  qui résout l'équation (18.767).

Pour être rigoureux, nous devons aussi montrer que la fonction  $\theta(t)$  est continue. Pour cela il faudrait utiliser le théorème de la fonction implicite 17.49. Nous verrons dans l'exemple 18.244 comment s'en sortir sans théorème de la fonction implicite, au prix de plus de calculs.  $\triangle$

Les coordonnées polaires sont données par le difféomorphisme

$$\begin{aligned} g: ]0, \infty[ \times ]0, 2\pi[ &\rightarrow \mathbb{R}^2 \setminus D \\ (r, \theta) &\mapsto (r \cos(\theta), r \sin(\theta)) \end{aligned} \quad (18.768)$$

où  $D$  est la demi-droite  $y = 0, x \geq 0$ . Le fait que les coordonnées polaires ne soient pas un difféomorphisme sur tout  $\mathbb{R}^2$  n'est pas un problème pour l'intégration parce que le manque de difféomorphisme est de mesure nulle dans  $\mathbb{R}^2$ . Le jacobien est donné par

$$Jg = \det \begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix} = r. \quad (18.769)$$

La fonction qui donne les coordonnées polaires est

$$\begin{aligned} \varphi: \mathbb{R}^+ \times ]0, 2\pi[ &\rightarrow \mathbb{R}^2 \\ (r, \theta) &\mapsto \begin{pmatrix} r \cos(\theta) \\ r \sin(\theta) \end{pmatrix}. \end{aligned} \quad (18.770)$$

Son Jacobien vaut

$$J_\varphi(r, \theta) = \det \begin{pmatrix} \frac{\partial x(r, \theta)}{\partial r} & \frac{\partial x(r, \theta)}{\partial \theta} \\ \frac{\partial y(r, \theta)}{\partial r} & \frac{\partial y(r, \theta)}{\partial \theta} \end{pmatrix} = \begin{vmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{vmatrix} = r. \quad (18.771)$$

### Proposition 18.229.

Soit la fonction

$$\begin{aligned} T: ]0, +\infty[ \times \mathbb{R} &\rightarrow \mathbb{R}^2 \setminus \{(0, 0)\} \\ (r, \theta) &\mapsto (r \cos(\theta), r \sin(\theta)). \end{aligned} \quad (18.772)$$

- (1) Elle est surjective.
- (2) Pour tout  $a \in \mathbb{R}$ , l'application  $T$  est bijective sur la bande  $]0, +\infty[ \times [a - \pi, a + \pi[$ .
- (3) Si  $a = 0$ , la fonction inverse est donnée par

$$T^{-1}(x, y) = (\sqrt{x^2 + y^2}, \arctan(y/x)). \quad (18.773)$$

Soit  $P = (x, y)$  un élément dans  $\mathbb{R}^2$ , on dit que  $r = \sqrt{x^2 + y^2}$  est le rayon de  $P$  et que  $\theta = \arctan(y/x)$  est son argument principal. L'origine ne peut pas être décrite en coordonnées polaires parce que si son rayon est manifestement zéro, on ne peut pas lui associer une valeur univoque de l'angle  $\theta$ .

### Exemple 18.230.

L'équation du cercle de rayon  $a$  et centre  $(0, 0)$  en coordonnées polaires est  $r = a$ .  $\triangle$

### Exemple 18.231.

Une équation possible pour la demi-droite  $x = y, x > 0$ , est  $\theta = \pi/4$ .  $\triangle$

## 18.13.1.3 Polaires : calcul de limites

EXooBEROooPhPsSU

**Exemple 18.232.**

Calculer les limites suivantes :

$$(1) \lim_{(x,y) \rightarrow (0,0)} \frac{(xy)^2}{(x+y)^2 + (x-y)^2}$$

$$(2) \lim_{(x,y) \rightarrow (0,0)} \frac{xy^3}{x^2 + y^2}$$

$$(3) \lim_{(x,y) \rightarrow (0,0)} \frac{x \sin(y)}{\sqrt{x^2 + y^2}}$$

Tentez de les faire par vous-même avant de regarder la solution qui suit.

- (1) Si  $(x, y) \in B(0, \delta) \setminus \{(0, 0)\}$ , alors le théorème 18.226 dit qu'il existe  $0 < r < \delta$  et  $\theta \in [0, 2\pi[$  tels que  $x = r \cos(\theta)$  et  $y = r \sin(\theta)$ . Avec ça nous avons

$$f(x, y) = \frac{r^4 \cos^2(\theta) \sin^2(\theta)}{r^2 (\cos(\theta) + \sin(\theta))^2 + r^2 (\cos(\theta) - \sin(\theta))^2}. \quad (18.774)$$

Vu que  $r \neq 0$  nous pouvons simplifier par  $r^2$ . Ce qui reste au dénominateur est une fonction de  $\theta$  continue sur le compact  $[0, 2\pi]$ . Je ne peux pas vous dire quel est le minimum, mais il est facile de voir que ce n'est pas zéro. Bref,

$$f(x, y) = r^2 f(\theta) \quad (18.775)$$

et  $|f(x, y)| < ar^2$  où  $a$  est une constante que vous pouvez vous amuser à calculer. En posant  $\delta < \sqrt{\epsilon/a}$  nous avons  $|f(x, y)| < \epsilon$  pour tout  $(x, y) \in B(0, \delta)$ , ce qui signifie que  $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$ .

- (2) Regardons la technique des coordonnées polaires. Nous remplaçons  $x$  par  $r \cos(\theta)$  et  $y$  par  $r \sin(\theta)$  :

$$f(r, \theta) = \frac{r^4 \cos(\theta) \sin^3(\theta)}{r^2} = r^2 \cos(\theta) \sin^3(\theta). \quad (18.776)$$

Cette fonction tend vers zéro quand  $r \rightarrow 0$ . Nous avons donc

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0. \quad (18.777)$$

Pour cet exercice nous pouvons aussi utiliser la règle de l'étau en écrivant d'abord

$$0 \leq |f(x, y)| \leq \frac{|x||y|^3}{|x^2 + y^2|}. \quad (18.778)$$

Mais on a  $|x| \leq \sqrt{x^2 + y^2}$ ,  $|y| \leq \sqrt{x^2 + y^2}$  et  $|x^2 + y^2| = (\sqrt{x^2 + y^2})^2$ , donc

$$0 \leq |f(x, y)| \leq \frac{\sqrt{x^2 + y^2} (\sqrt{x^2 + y^2})^3}{(\sqrt{x^2 + y^2})^2} = (\sqrt{x^2 + y^2})^2 \rightarrow 0. \quad (18.779)$$

- (3) En passant aux polaires, nous avons

$$f(r, \theta) = \frac{r \cos \theta \sin(r \sin \theta)}{r} = \cos(\theta) \sin(r \sin \theta). \quad (18.780)$$

La limite de cette dernière fonction lorsque  $r \rightarrow 0$  vaut zéro.

Une autre façon de procéder consiste à multiplier et diviser par  $y$  de telle façon à faire apparaître  $\sin(y)/y$  dont nous connaissons la limite :

$$f(x, y) = \frac{\sin(y)}{y} \cdot \frac{xy}{\sqrt{x^2 + y^2}}. \quad (18.781)$$

La limite du premier facteur est 1, tandis que le second peut être traité de façon classique en prenant la valeur absolue et en majorant  $|x|$  par  $\sqrt{x^2 + y^2}$ .

△

**18.13.1.4 Transformation inverse : théorie**

Voyons la question inverse : comment retrouver  $r$  et  $\theta$  si on connaît  $x$  et  $y$  ? Tout d'abord,

$$r = \sqrt{x^2 + y^2} \quad (18.782)$$

parce que la coordonnée  $r$  est la distance entre l'origine et  $(x, y)$ . Comment trouver l'angle ? Nous supposons  $(x, y) \neq (0, 0)$ . Si  $x = 0$ , alors le point est sur l'axe vertical et nous avons

$$\theta = \begin{cases} \pi/2 & \text{si } y > 0 \\ 3\pi/2 & \text{si } y < 0 \end{cases} \quad (18.783)$$

Notez que si  $y < 0$ , conformément à notre convention  $\theta \geq 0$ , nous avons noté  $\frac{3\pi}{2}$  et non  $-\frac{\pi}{2}$ .

Supposons maintenant le cas général avec  $x \neq 0$ . Les équations (18.748) montrent que

$$\tan(\theta) = \frac{y}{x}. \quad (18.784)$$

Nous avons donc

$$\theta = \tan^{-1}\left(\frac{y}{x}\right). \quad (18.785)$$

La fonction inverse de la fonction tangente est celle définie plus haut.

**18.13.1.5 Transformation inverse : pratique**

Le code suivant utilise Sage.

```

1  #! /usr/bin/sage -python
2  # -*- coding: utf8 -*-
3
4  from sage.all import *
5
6  def PointToPolaire(x,y):
7      x=SR(x)
8      y=SR(y)
9      r = sqrt(x**2+y**2)
10     if x == 0:
11         if y > 0:
12             alpha = pi/2
13         if y < 0:
14             alpha = 3*pi/2
15         if y == 0 :
16             raise ValueError,"Pas d'angle pour le point (0,0) !!"
17     else :
18         alpha = atan(y/x)
19     if (x < 0) and (y == 0) :
20         alpha = pi
21     if (x < 0) and (y > 0) :
22         alpha = alpha + pi
23     if (x < 0) and (y < 0) :
24         alpha = alpha + pi
25     alpha=alpha.simplify_trig()
26     return (r,alpha)
27
28 print PointToPolaire(1,1)
29 print PointToPolaire(-2,1)

```

```
30 | print PointToPolaire(6*sqrt(3)/2,3)
```

tex/frido/calculAngle.py

Son exécution retourne :

```
(sqrt(2), 1/4*pi)
(sqrt(5), pi - arctan(1/2))
(6, 1/6*pi)
```

Notez que ce sont des valeurs *exactes*. Ce ne sont pas des approximations, Sage travaille de façon symbolique.

### 18.13.1.6 Coordonnées polaires : dérivées partielles

Le changement de coordonnées pour les coordonnées polaires est la fonction

$$f \begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}. \quad (18.786)$$

Considérons une fonction  $g$  sur  $\mathbb{R}^2$ , et définissons la fonction  $\tilde{g}$  par

$$\tilde{g}(r, \theta) = g(r \cos \theta, r \sin \theta). \quad (18.787)$$

La formule (12.817) permet de trouver les dérivées partielles de  $g$  par rapport à  $r$  et  $\theta$  en termes de celles par rapport à  $x$  et  $y$  de  $g$ .

Pour faire le lien avec les notations du point précédent, nous avons

$$\begin{aligned} f_1(r, \theta) &= r \cos(\theta) \\ f_2(r, \theta) &= r \sin(\theta) \\ (x_1, x_2) &\rightarrow (r, \theta) \\ (y_1, y_2) &\rightarrow (x, y). \end{aligned} \quad (18.788)$$

Nous avons donc

$$\begin{aligned} \frac{\partial \tilde{g}}{\partial r}(r, \theta) &= \sum_{i=1}^2 \frac{\partial g}{\partial x_i}(f(r, \theta)) \frac{\partial f_i}{\partial r}(r, \theta) \\ &= \frac{\partial g}{\partial x}(r \cos \theta, r \sin \theta) \frac{\partial (r \cos \theta)}{\partial r}(r, \theta) \\ &\quad + \frac{\partial g}{\partial y}(r \cos \theta, r \sin \theta) \frac{\partial (r \sin \theta)}{\partial r}(r, \theta) \\ &= \cos \theta \frac{\partial g}{\partial x}(r \cos \theta, r \sin \theta) + \sin \theta \frac{\partial g}{\partial y}(r \cos \theta, r \sin \theta). \end{aligned} \quad (18.789)$$

Prenons par exemple  $g(x, y) = \frac{1}{x^2 + y^2}$ . Étant donné que

$$\frac{\partial g}{\partial x} = \frac{-2x}{(x^2 + y^2)^2}, \quad (18.790)$$

nous avons

$$\frac{\partial g}{\partial x}(r \cos \theta, r \sin \theta) = \frac{-2 \cos \theta}{r^3}. \quad (18.791)$$

En utilisant la formule,

$$\frac{\partial \tilde{g}}{\partial r}(r, \theta) = \cos(\theta) \left( \frac{-2 \cos \theta}{r^3} \right) + \sin(\theta) \left( \frac{-2 \sin \theta}{r^3} \right) = -\frac{2}{r^3}. \quad (18.792)$$

Nous pouvons vérifier directement que cela est correct. En effet

$$\tilde{g}(r, \theta) = g(r \cos \theta, r \sin \theta) = \frac{1}{r^2}, \quad (18.793)$$

dont la dérivée par rapport à  $r$  vaut  $-2/r^3$ .

En ce qui concerne la dérivée par rapport à  $\theta$ , nous avons

$$\begin{aligned} \frac{\partial \tilde{g}}{\partial \theta} &= \frac{\partial g}{\partial x}(r \cos \theta, r \sin \theta) \frac{\partial(r \cos(\theta))}{\partial \theta} + \frac{\partial g}{\partial y}(r \cos \theta, r \sin \theta) \frac{\partial(r \sin(\theta))}{\partial \theta} \\ &= \left( \frac{-2 \cos \theta}{r^3} \right) (-r \sin \theta) + \left( \frac{-2 \sin \theta}{r^3} \right) (r \cos \theta) \\ &= 0. \end{aligned} \quad (18.794)$$

En résumé et avec quelques abus de notation :

$$\begin{aligned} \frac{\partial \tilde{g}}{\partial r} &= \cos(\theta) \frac{\partial g}{\partial x} + \sin(\theta) \frac{\partial g}{\partial y} \\ \frac{\partial \tilde{g}}{\partial \theta} &= -r \sin(\theta) \frac{\partial g}{\partial x} + r \cos(\theta) \frac{\partial g}{\partial y} \end{aligned} \quad (18.795)$$

### 18.13.2 Coordonnées cylindriques

Les **coordonnées cylindriques** sont un perfectionnement des coordonnées polaires. Il s'agit simplement de donner le point  $(x, y, z)$  en faisant la conversion  $(x, y) \mapsto (r, \theta)$  et en gardant le  $z$ . Les formules de passage sont

$$\begin{cases} x = r \cos(\theta) & (18.796a) \\ y = r \sin(\theta) & (18.796b) \\ z = z. & (18.796c) \end{cases}$$

Soit  $T$  la fonction de  $]0, +\infty[ \times \mathbb{R}^2$  dans  $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$  définie par

$$\begin{aligned} T : \quad ]0, +\infty[ \times \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{R}^3 \setminus \{(0, 0, 0)\} \\ (r, \theta, z) &\mapsto (r \cos \theta, r \sin \theta, z), \end{aligned} \quad (18.797)$$

Cette fonction est surjective. Elle est bijective sur chaque bande de la forme  $]0, +\infty[ \times [a - \pi, a + \pi[ \times \mathbb{R}$ ,  $a$  dans  $\mathbb{R}$ . Il n'y a presque rien de nouveau par rapport aux coordonnées polaires. Les coordonnées cylindriques sont intéressantes si on décrit un objet invariant par rapport aux rotations autour de l'axe des  $z$ .

#### Exemple 18.233.

Il faut savoir ce que décrivent les équations les plus simples en coordonnées cylindriques,

- $r \leq a$ , pour  $a$  constant dans  $]0, +\infty[$ , est le cylindre de hauteur infinie qui a pour axe l'axe des  $z$  et pour base le disque de rayon  $a$  centré à l'origine,
- $r = a$  est la surface du cylindre,
- $\theta = b$  est un demi-plan ouvert et sa fermeture contient l'axe des  $z$ ,
- $z = c$  est un plan parallèle au plan  $xy$ .

△

#### Exemple 18.234.

Un demi-cône qui a son sommet en l'origine et pour axe l'axe des  $z$  est décrit par  $z = dr$ . Si  $d$  est positif il s'agit de la moitié supérieure du cône, si  $d < 0$  de la moitié inférieure. △

#### Exemple 18.235.

De même, la sphère de rayon  $a$  et centrée à l'origine est l'assemblage des calottes  $z = \sqrt{a^2 - r^2}$  et  $z = -\sqrt{a^2 - r^2}$ . △

En ce qui concerne les coordonnées cylindriques, le Jacobien est donné par

$$J(r, \theta, z) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r. \quad (18.798)$$

Nous avons donc  $dx dy dz = r dr d\theta dz$ .

$$\begin{cases} x = r \cos \theta & (18.799a) \\ y = r \sin \theta & (18.799b) \\ z = z & (18.799c) \end{cases}$$

avec  $r \in ]0, \infty[$ ,  $\theta \in [0, 2\pi[$  et  $z \in \mathbb{R}$ . Le jacobien vaut  $r$ .

### 18.13.3 Coordonnées sphériques

Soit  $T$  la fonction de  $]0, +\infty[ \times \mathbb{R}^2$  dans  $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$  définie par

$$\begin{aligned} T : ]0, +\infty[ \times \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{R}^3 \setminus \{(0, 0, 0)\} \\ (\rho, \theta, \phi) &\mapsto (\rho \cos \theta \sin \phi, \rho \sin \theta \sin \phi, \rho \cos \phi), \end{aligned} \quad (18.800)$$

Cette fonction est surjective. Elle est bijective sur chaque bande de la forme  $]0, +\infty[ \times [a - \pi, a + \pi[ \times [b - \pi/2, b + \pi/2[$ ,  $a$  et  $b$  dans  $\mathbb{R}$ . Si  $a = 0$  et  $b = -\pi/2$  la fonction inverse  $T^{-1}$  est donnée par

$$\begin{aligned} T^{-1} : \mathbb{R}^3 \setminus \{(0, 0, 0)\} &\rightarrow ]0, +\infty[ \times [-\pi, \pi[ \times [0, \pi[ \\ (x, y, z) &\mapsto \left( \sqrt{x^2 + y^2 + z^2}, \arctan \frac{y}{x}, \arccos \left( \frac{z}{\sqrt{x^2 + y^2 + z^2}} \right) \right). \end{aligned} \quad (18.801)$$

Soit  $P$  un point dans  $\mathbb{R}^3$ . L'angle  $\phi$  est l'angle entre le demi-axe positif des  $z$  et le vecteur  $\overrightarrow{OP}$ ,  $\rho$  est la norme de  $\overrightarrow{OP}$  et  $\theta$  est l'argument en coordonnées polaires de la projection de  $\overrightarrow{OP}$  sur le plan  $x$ - $y$ .

#### Remarque 18.236.

Dans la littérature, les angles  $\theta$  et  $\phi$  sont parfois inversés (voire, changent de nom, par exemple  $\varphi$  au lieu de  $\phi$ ). Il faut donc être très prudent lorsqu'on veut utiliser dans un cours des formules données dans un autre cours.

#### Exemple 18.237.

Il faut connaître le sens des équations plus simples,

- $\rho \leq a$ , pour  $a$  constant dans  $]0, +\infty[$ , est la boule fermée de rayon  $a$  centrée à l'origine,
- $\rho = a$  est la sphère de rayon  $a$  centrée à l'origine,
- $\theta = b$  est un demi-plan ouvert et sa fermeture contient l'axe des  $z$ ,
- $\phi = c$  est un demi-cône qui a son sommet à l'origine et pour axe l'axe des  $z$ . Si  $c$  est positif il s'agit de la moitié supérieure du cône, si  $d < 0$  de la moitié inférieure.

△

Les **coordonnées sphériques** sont ce qu'on appelle les « méridiens » et « longitudes » en géographie. Les formules de transformation sont l

$$\begin{cases} x = \rho \sin(\theta) \cos(\varphi) & (18.802a) \\ y = \rho \sin(\theta) \sin(\varphi) & (18.802b) \\ z = \rho \cos(\theta) & (18.802c) \end{cases}$$

avec  $0 \leq \theta \leq \pi$  et  $0 \leq \varphi < 2\pi$ .



**Remarque 18.238.**

Attention : d'un livre à l'autre les conventions sur les noms des angles changent. N'essayez donc pas d'étudier par cœur des formules concernant les coordonnées sphériques trouvées autre part. Par exemple sur le premier dessin de [wikipédia](#), l'angle  $\varphi$  est noté  $\theta$  et l'angle  $\theta$  est noté  $\Phi$ . Mais vous noterez que sur cette même page, les conventions de noms de ces angles changent plusieurs fois.

Les coordonnées sphériques sont données par

$$\begin{cases} x = r \cos \theta \sin \varphi \\ y = r \sin \theta \sin \varphi \\ z = r \cos \varphi \end{cases} \quad \text{EqChmVarSpherique (18.803)}$$

avec  $r \in ]0, \infty[$ ,  $\theta \in ]0, 2\pi[$  et  $\varphi \in ]0, \pi[$ . Le jacobien associé est  $Jg(r, \theta, \varphi) = -r^2 \sin \varphi$ . N'oubliez pas que lorsqu'on effectue un changement de variables dans une intégrale, la *valeur absolue* du jacobien apparaît.

Cependant notre convention de coordonnées sphériques fait venir  $\sin(\phi)$  avec  $\phi \in [0, \pi[$ ; vu que le signe de  $\sin(\phi)$  y est toujours positif, cette histoire de valeur absolue est sans grandes conséquences. Ce n'est pas le cas de toutes les conventions possibles.

**18.13.3.1 Coordonnées sphériques : inverse**

Trouvons le changement inverse, c'est-à-dire trouvons  $\rho$ ,  $\theta$  et  $\varphi$  en termes de  $x$ ,  $y$  et  $z$ . D'abord nous avons

$$\rho = \sqrt{x^2 + y^2 + z^2}. \quad (18.804)$$

Ensuite nous savons que

$$\cos(\theta) = \frac{z}{\rho} \quad (18.805)$$

détermine de façon unique<sup>87</sup> un angle  $\theta \in [0, \pi]$ . Dès que  $\rho$  et  $\theta$  sont connus, nous pouvons poser  $r = \rho \sin \theta$  et alors nous nous trouvons avec les équations

$$\begin{cases} x = r \cos(\varphi) \\ y = r \sin(\varphi), \end{cases} \quad \begin{matrix} (18.806a) \\ (18.806b) \end{matrix}$$

qui sont similaires à celles déjà étudiées dans le cas des coordonnées polaires.

**18.14 Calcul de limites**

Beaucoup de techniques de calcul de limites fonctionnent bien avec les fonctions trigonométriques, entre autres grâce à l'utilisation des coordonnées polaires de la proposition 18.229. Nous en voyons quelques exemples à présent.

**Exemple 18.239** (Limite et prolongement par continuité).

La fonction

$$f(x) = \frac{\cos(x) - 1}{x} \quad (18.807)$$

n'est pas définie en  $x = 0$ .

Nous avons vu dans l'équation (18.6) que  $\cos(0) = 1$ , donc la limite

$$\lim_{x \rightarrow 0} \frac{\cos(x) - 1}{x} \quad (18.808)$$

est la limite définissant la dérivée de cosinus en 0 (ici, le  $x$  joue le rôle de  $\epsilon$ ). Le lemme 18.3 nous donne la dérivée du cosinus comme étant le sinus. Nous avons donc :

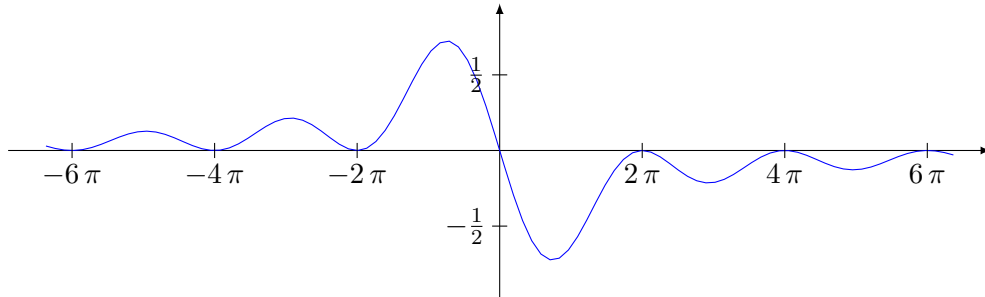
$$\lim_{x \rightarrow 0} \frac{\cos(x) - 1}{x} = \sin(0) = 0, \quad (18.809)$$

87. Le problème  $\rho = 0$  ne se pose pas ; pourquoi ?

et nous définissons le prolongement par continuité :

$$\tilde{f}(x) = \begin{cases} \frac{\cos(x)-1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases} \quad (18.810)$$

Encore une fois, le graphe de la fonction  $\tilde{f}$  ne présente aucune particularité autour de  $x = 0$ .



△

EXooINLRooPzRWEA

**Exemple 18.240** (Un calcul heuristique de limite).

Soit à calculer la limite suivante :

$$\lim_{x \rightarrow 0} \frac{e^{-2 \cos(x)+2} \sin(x)}{\sqrt{e^{2 \cos(x)+2} - 1}}. \quad (18.811)$$

La stratégie que nous allons suivre pour calculer cette limite est de développer certaines parties de l'expression en série de Taylor, afin de simplifier l'expression. La première chose à faire est de remplacer  $e^{y(x)}$  par  $1 + y(x)$  lorsque  $y(x) \rightarrow 0$ . La limite devient

$$\lim_{x \rightarrow 0} \frac{(-2 \cos(x) + 3) \sin(x)}{\sqrt{-2 \cos(x) + 2}}. \quad (18.812)$$

Nous allons maintenant remplacer  $\cos(x)$  par  $1 - x^2/2$  au numérateur et par  $1 - x^2/2$  au dénominateur. Pourquoi ? Parce que le cosinus du dénominateur est dans une racine, donc nous nous attendons à ce que le terme de degré deux du cosinus donne un degré un en dehors de la racine, alors que du degré un est exactement ce que nous avons au numérateur : le développement du sinus commence par  $x$ .

Nous calculons donc

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{\sqrt{-2 \left(1 - \frac{x^2}{2}\right) + 2}} = \lim_{x \rightarrow 0} \frac{\sin(x)}{x} = 1. \quad (18.813)$$

Tout ceci n'est évidemment pas très rigoureux, mais en principe vous avez tous les éléments en main pour justifier les étapes. △

### 18.14.1 Méthode des coordonnées polaires

SUBSECooWCGMooPrXSpt

La proposition suivante exprime la définition de la limite en d'autres termes, et va être pratique dans le calcul de certaines limites.

PropMethodePolaire

**Proposition 18.241.**

Soit  $f: D \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ ,  $a$  un point d'accumulation de  $D$  et  $\ell \in \mathbb{R}^n$ . Nous définissons

$$E_r = \{f(x) \text{ tel que } x \in B(a, r) \cap D\}, \quad (18.814)$$

et

$$s_r = \sup\{\|v - \ell\| \text{ tel que } v \in E_r\}. \quad (18.815)$$

Alors nous avons  $\lim_{x \rightarrow a} f(x) = \ell$  si et seulement si  $\lim_{r \rightarrow 0} s_r = 0$ .

Dans cette proposition,  $E_r$  représente l'ensemble des valeurs atteintes par  $f$  dans un rayon  $r$  autour de  $a$ . Le nombre  $s_r$  sélectionne, parmi toutes ces valeurs, celle qui est la plus éloignée de  $\ell$  et donne la distance. En d'autres termes,  $s_r$  est la distance maximale entre  $f(x)$  et  $\ell$  lorsque  $x$  est à une distance au maximum  $r$  de  $a$ .

Lorsque nous avons affaire à une fonction  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , cette proposition nous permet de calculer facilement les limites en passant aux coordonnées polaires.

ExempleMethodeTrigigi

**Exemple 18.242.**

Reprenons la fonction de l'exemple 12.243 :

$$f(x, y) = \frac{xy}{x^2 + y^2}. \quad (18.816)$$

Son domaine est  $\mathbb{R}^2 \setminus \{(0, 0)\}$ . Nous voulons calculer  $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$ . Écrivons la définition de  $E_r$  :

$$E_r = \{f(x, y) \text{ tel que } (x, y) \in B((0, 0), r)\}. \quad (18.817)$$

Les points de la boule sont, en coordonnées polaires, les points de la forme  $(\rho, \theta)$  avec  $\rho < r$ . La chose intéressante est que  $f(\rho, \theta)$  est relativement simple (plus simple que la fonction départ). En effet en remplaçant tous les  $x$  par  $\rho \cos(\theta)$  et tous les  $y$  par  $\rho \sin(\theta)$ , et en utilisant le fait que  $\cos^2(\theta) + \sin^2(\theta) = 1$ , nous trouvons

$$f(\rho, \theta) = \frac{\rho^2 \cos(\theta) \sin(\theta)}{\rho^2} = \cos(\theta) \sin(\theta). \quad (18.818)$$

Eq2807fpolairerhodeuxcossin

Cela signifie que

$$E_r = \{\cos(\theta) \sin(\theta) \text{ tel que } \theta \in [0, 2\pi[ \}. \quad (18.819)$$

Prenons  $\ell$  quelconque. Le nombre  $s_r$  est le supremum des

$$\|\ell - \cos(\theta) \sin(\theta)\| \quad (18.820)$$

lorsque  $\theta$  parcourt  $[0, 2\pi]$ . Nous ne sommes pas obligés de calculer la valeur exacte de  $s_r$ . Ce qui compte ici est que  $s_r$  ne vaut certainement pas zéro, et ne dépend pas de  $r$ . Donc il est impossible d'avoir  $\lim_{r \rightarrow 0} s_r = 0$ , et la fonction donnée n'a pas de limite en  $(0, 0)$ .  $\triangle$

Nous pouvons retenir cette règle pour calculer les limites lorsque  $(x, y) \rightarrow (0, 0)$  de fonctions  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  :

- (1) passer en coordonnées polaires, c'est-à-dire remplacer  $x$  par  $\rho \cos(\theta)$  et  $y$  par  $\rho \sin(\theta)$  ;
- (2) nous obtenons une fonction  $g$  de  $\rho$  et  $\theta$ . Si la limite  $\lim_{r \rightarrow 0} g(r, \theta)$  n'existe pas ou dépend de  $\theta$ , alors la fonction n'a pas de limite. Si on peut majorer  $g$  par une fonction ne dépendant pas de  $\theta$ , et que cette fonction a une limite lorsque  $r \rightarrow 0$ , alors cette limite est la limite de la fonction.

La vraie difficulté de la technique des coordonnées polaires est de trouver le supremum de  $E_r$ , ou tout au moins de montrer qu'il est borné par une fonction qui a une limite qui ne dépend pas de  $\theta$ . Une des situations classiques dans laquelle c'est facile est lorsque la fonction se présente comme une fonction de  $r$  multiplié par une fonction de  $\theta$ .

Exemplexyxsqysq

**Exemple 18.243.**

Soit à calculer la limite

$$\lim_{(x,y) \rightarrow (0,0)} xy \left( \frac{x^2 - y^2}{x^2 + y^2} \right). \quad (18.821)$$

Le passage aux coordonnées polaires donne

$$f(r, \theta) = r^2 \sin \theta \cos \theta (\cos^2 \theta - \sin^2 \theta). \quad (18.822)$$

Déterminer le supremum de cela est relativement difficile. Mais nous savons que de toutes façons, la quantité  $\sin \theta \cos \theta (\cos^2 \theta - \sin^2 \theta)$  est bornée par 1. Donc

$$\|f(r, \theta)\| \leq r^2. \quad (18.823)$$

Maintenant la règle de l'étau montre que  $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$  est zéro.

La situation vraiment gênante serait celle avec une fonction de  $\theta$  qui risque de s'annuler dans un dénominateur.  $\triangle$

L'exemple 18.228 donnera un cas où la méthode fonctionne plus difficilement. Entre autres parce qu'il utilisera en même temps la méthode des chemins et celle des coordonnées polaires.

ExmeASDLAf

### Exemple 18.244.

Considérons fonction

$$f(x, y) = \frac{x^2 + y^2}{x - y}. \quad (18.824)$$

Une mauvaise idée pour prouver que la limite n'existe pas pour  $(x, y) \rightarrow (0, 0)$  est de considérer le chemin  $(t, t)$ . En effet, la fonction n'existe pas sur ce chemin. Or la méthode des chemins parle uniquement de chemins contenus dans le domaine de la fonction.

Nous prouvons que la limite n'existe pas en trouvant des chemins le long desquels les limites sont différentes. Si nous essayons le chemin  $(t, kt)$  avec  $k$  constant, nous trouvons

$$f(t, kt) = \frac{t(1 + k^2)}{1 - k}. \quad (18.825)$$

La limite  $t \rightarrow 0$  est hélas toujours 0. Nous ne pouvons donc pas conclure.

Nous allons maintenant utiliser la même technique que celle utilisée en coordonnées polaires. Vous noterez que dans ce cas, travailler en cartésiennes donne lieu à des calculs plus longs. L'astuce consiste à prendre  $k$  non constant et à chercher par exemple  $k(t)$  de façon à avoir

$$\frac{1 + k(t)^2}{1 - k(t)} = \frac{1}{t}. \quad (18.826)$$

Avec une telle fonction  $k$ , la fonction  $t \mapsto f(t, tk(t))$  serait la constante 1. L'équation à résoudre pour  $k$  est

$$tk^2 + k + (t - 1) = 0, \quad (18.827)$$

et les solutions sont

$$k(t) = \frac{-1 \pm \sqrt{1 - 4t(t - 1)}}{2t}. \quad (18.828)$$

Nous proposons donc les chemins

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ \frac{-1 \pm \sqrt{1 - 4t(t - 1)}}{2} \end{pmatrix} \quad (18.829)$$

Nous devons vérifier deux points. D'abord que ce chemin est bien défini, et ensuite que  $tk(t)$  tend bien vers zéro lorsque  $t \rightarrow 0$  (sinon  $(t, k(t)t)$  n'est pas un chemin passant par  $(0, 0)$ ). Lorsque  $t$  est petit, ce qui se trouve sous la racine est proche de 1 et ne pose pas de problèmes. Ensuite,

$$\lim_{t \rightarrow 0} tk(t) = \frac{-1 \pm 1}{2}. \quad (18.830)$$

En choisissant le signe +, nous trouvons un chemin qui nous convient.

Ce que nous avons prouvé est que

$$f\left(t, \frac{-1 + \sqrt{1 - 4t(t - 1)}}{2}\right) = 1 \quad (18.831)$$

pour tout  $t$ . Le long de ce chemin, la limite de  $f$  est donc 1. Cette limite est différente des limites obtenues le long de chemins avec  $k$  constant. La limite  $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$  n'existe donc pas.  $\triangle$

**Exemple 18.245.**

Considérons la fonction (figure 18.11)

$$f(x, y) = \begin{cases} \sqrt{x^2 + y^2} \sin \frac{1}{\sqrt{x^2 + y^2}} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0), \end{cases} \quad (18.832)$$

et cherchons la limite  $(x, y) \rightarrow (0, 0)$ . Le passage en coordonnées polaires<sup>88</sup> donne

$$f(\rho, \theta) = \rho \sin \frac{1}{\rho}. \quad (18.833)$$

Pour calculer la limite de cela lorsque  $\rho \rightarrow 0$ , nous remarquons que

$$0 \leq \left| \rho \sin \frac{1}{\rho} \right| \leq \rho \quad (18.834)$$

parce que  $\sin(\frac{1}{\rho}) \leq 1$  quel que soit  $\rho$ . Or évidemment  $\lim_{\rho \rightarrow 0} \rho = 0$ , donc la limite de la fonction (18.833) est zéro et ne dépend pas de  $\theta$ . Nous en concluons que  $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$ .  $\triangle$

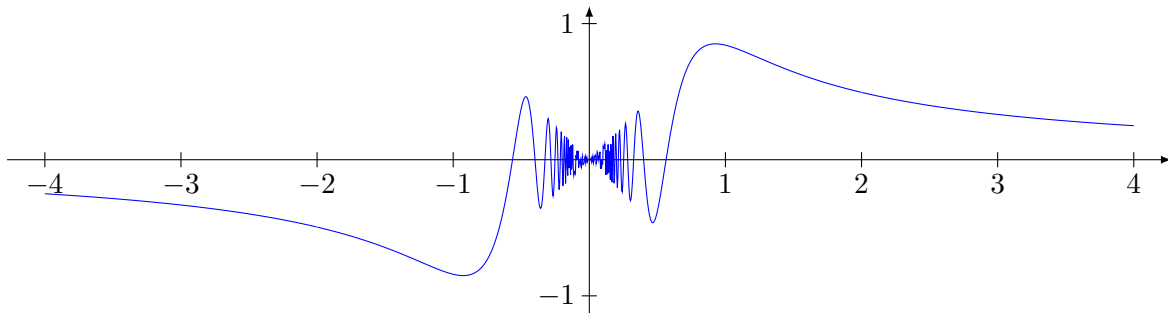


FIGURE 18.11: La fonction de l'exemple 18.245.

### 18.14.2 Méthode du développement asymptotique

Nous savons que nous pouvons développer certaines fonctions en série grâce au développement de Taylor (théorème 12.498). Lorsque nous avons une limite à calculer, nous pouvons remplacer certaines parties de la fonction à traiter par la formule (12.1406a). Cela est très utile pour comparer des fonctions trigonométrique à des polynômes.

**Lemme 18.246.**

Nous avons la limite

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = 1. \quad (18.835)$$

*Démonstration.* Une manière de prouver cela est d'écrire

$$\sin(x) = x + h(x) \quad (18.836)$$

avec  $h \in o(x)$ , c'est-à-dire  $\lim_{x \rightarrow 0} h(x)/x = 0$ . Alors nous avons

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{x + h(x)}{x} = \lim_{x \rightarrow 0} \frac{x}{x} + \lim_{x \rightarrow 0} \frac{h(x)}{x} = 1. \quad (18.837)$$

□

L'utilisation de la proposition 12.534 permet d'utiliser cette technique dans le cadre de limites à plusieurs variables. Reprenons le lemme 18.246 un tout petit peu modifié :

<sup>88</sup>. Proposition 18.229.

LEMooSFALooVRbNb

**Lemme 18.247.**

Pour tout  $x > 0$  nous avons  $\sin(x) < x$ .

*Démonstration.* Nous posons  $f(x) = x - \sin(x)$ . Cette fonction vérifie  $f(0) = 0$  et

$$f'(x) = 1 - \cos(x). \quad (18.838)$$

Vu que  $|\cos(x)| \leq 1$ , nous avons toujours  $f'(x) \geq 0$  et même  $f'(x) > 0$  pour  $x \in ]0, \delta]$ . Donc  $f$  est au moins strictement croissante sur  $]0, \delta]$  et ensuite strictement croissante presque partout.  $\square$

**Exemple 18.248.**

Soit à calculer  $\lim_{(x,y) \rightarrow (0,0)} f(x,y)$  où

$$f(x,y) = \frac{\sin(xy)}{xy}. \quad (18.839)$$

La première chose à faire est de voir  $f$  comme la composée de fonctions  $f = f_1 \circ f_2$  avec

$$\begin{aligned} f_1: \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto \frac{\sin(t)}{t} \end{aligned} \quad (18.840)$$

et

$$\begin{aligned} f_2: \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x,y) &\mapsto xy. \end{aligned} \quad (18.841)$$

Étant donné que  $\lim_{(x,y) \rightarrow (0,0)} f_2(x,y) = 0$ , nous avons  $\lim_{(x,y) \rightarrow (0,0)} f(x,y) = \lim_{t \rightarrow 0} f_1(t) = 1$ .  $\triangle$

EXooETZYooYsKPDJ

**Exemple 18.249.**

Les dérivées partielles de la fonction  $f(x,y) = xy^3 + \sin y$  au point  $(0, \pi)$  sont

$$\begin{aligned} \partial_x f(0, \pi) &= \frac{\partial f}{\partial x}(0, \pi) = \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{(t\pi^3 + \sin \pi) - (\sin \pi)}{t} = \pi^3, \\ \partial_y f(0, \pi) &= \frac{\partial f}{\partial y}(0, \pi) = \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{0(\pi + t)^3 + \sin(t + \pi) - 0 \cdot \pi^3}{t} = \cos \pi = -1, \end{aligned}$$

 $\triangle$ 

## 18.15 Quelques intégrales avec de la trigonométrie

SECooOOPPooZLbaEH

Le théorème 14.282 manque un peu d'exemples. Nous allons en voir quelques-uns maintenant.

### 18.15.1 Changement de variables

Le domaine  $E = \{(x,y) \in \mathbb{R}^2 \text{ tel que } x^2 + y^2 < 1\}$  s'écrit plus facilement  $E = \{(r, \theta) \text{ tel que } r < 1\}$  en coordonnées polaires. Le passage aux coordonnées polaires permet de transformer une intégration sur un domaine rond à une intégration sur le domaine rectangulaire  $]0, 2\pi[ \times ]0, 1[$ . La question est évidemment de savoir si nous pouvons écrire

$$\int_E f = \int_0^{2\pi} \int_0^1 f(r \cos \theta, r \sin \theta) dr d\theta. \quad (18.842)$$

Hélas, non. La vie n'est pas aussi simple ; il faut tenir compte du fait que le changement de base dilate ou contracte certaines surfaces.

Soit  $\varphi: D_1 \subset \mathbb{R}^2 \rightarrow D_2 \subset \mathbb{R}^2$  une fonction bijective de classe  $C^1$  dont l'inverse est également de classe  $C^1$ . On désigne par  $x$  et  $y$  ses composantes, c'est-à-dire que

$$\varphi(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \end{pmatrix} \quad (18.843)$$

avec  $(u,v) \in D_1$ .

ThoChamDeVarIntDDf

**Théorème 18.250.**

Soit une fonction continue  $f: D_2 \rightarrow \mathbb{R}$ . Alors

$$\int_{\varphi(D_1)} f(x, y) dx dy = \int_{D_1} f(x(u, v), y(u, v)) |J_\varphi(u, v)| du dv \quad (18.844)$$

où  $J_\varphi$  est le Jacobien de  $\varphi$  c'est-à-dire

$$J_\varphi(u, v) = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}. \quad (18.845)$$

Ne pas oublier de prendre la valeur absolue lorsqu'on utilise le Jacobien dans un changement de variables.

**18.15.2 Coordonnées polaires****Exemple 18.251.**

Calculons la surface du disque  $D$  de rayon  $R$ . Nous devons calculer

$$\int_D dx dy. \quad (18.846)$$

Pour passer au polaires, nous savons que le disque est décrit par

$$D = \{(r, \theta) \text{ tel que } 0 \leq r \leq R, 0 \leq \theta \leq 2\pi\}. \quad (18.847)$$

Nous avons donc

$$\int_D dx dy = \int_D r dr d\theta = \int_0^{2\pi} \int_0^R r dr d\theta = 2\pi \int_0^R r dr = \pi R^2. \quad (18.848)$$

△

ExpnFDtAtV

**Exemple 18.252.**

Montrons comment intégrer la fonction  $f(x, y) = \sqrt{1 - x^2 - y^2}$  sur le domaine délimité par la droite  $y = x$  et le cercle  $x^2 + y^2 = y$ , représenté sur la figure 18.12. Pour trouver le centre et le rayon du cercle  $x^2 + y^2 = y$ , nous commençons par écrire  $x^2 + y^2 - y = 0$ , et ensuite nous reformons le carré :  $y^2 - y = (y - \frac{1}{2})^2 - \frac{1}{4}$ .

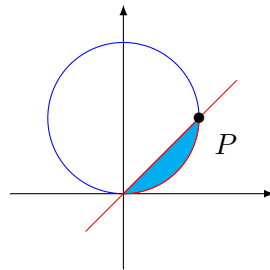


FIGURE 18.12: Passage en polaire pour intégrer sur un morceau de cercle.

ExpnFDtAtV

Le passage en polaire transforme les équations du bord du domaine en

$$\begin{aligned} \cos(\theta) &= \sin(\theta) \\ r^2 &= r \sin(\theta). \end{aligned} \quad (18.849)$$

L'angle  $\theta$  parcourt donc  $]0, \pi/4[$ , et le rayon, pour chacun de ces  $\theta$  parcourt  $]0, \sin(\theta)[$ . La fonction à intégrer se note maintenant  $f(r, \theta) = \sqrt{1 - r^2}$ . Donc l'intégrale à calculer est

$$\int_0^{\pi/4} \left( \int_0^{\sin(\theta)} \sqrt{1 - r^2} r dr \right). \quad (18.850)$$

PgOMRapIntMultFubiniBoutCercle

Remarquez la présence d'un  $r$  supplémentaire pour le jacobien.

Notez que les coordonnées du point  $P$  sont  $(1, 1)$ .

△

En pratique, lors du passage en coordonnées polaires, le «  $dx dy$  » devient «  $r dr d\theta$  ».

**Exemple 18.253.**

On veut évaluer l'intégrale de la fonction  $f(x, y) = x^2 + y^2$  sur la région  $V$  suivante :

$$V = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1, x > 0, y > 0\}.$$

On peut faire le calcul directement,

$$\int_V f(x, y) dV = \int_0^1 \int_0^{\sqrt{1-x^2}} x^2 + y^2 dy dx = \int_0^1 \left( x^2 \sqrt{1-x^2} + \frac{(1-x^2)^{3/2}}{3} \right) dx$$

mais c'est un peu ennuyeux. On peut simplifier beaucoup les calculs avec un changement de variables vers les coordonnées polaires. Dans ce cas, on sait bien que le difféomorphisme à utiliser est  $\phi(r, \theta) = (r \cos \theta, r \sin \theta)$ . Le jacobien  $J_\phi$  est

$$J_\phi(r, \theta) = \begin{vmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{vmatrix} = r, \quad (18.851)$$

qui est toujours positif. D'une part, la fonction  $f$  peut s'écrire sous la forme  $f(\phi(r, \theta)) = r^2$  et d'autre part,  $\phi^{-1}(V) = ]0, 1] \times ]0, \pi/2[$ . Par conséquent, la formule du changement de variables nous donne

$$\int_V f(x, y) dV = \int_0^{\pi/2} \int_0^1 r^3 dr d\theta = \int_0^{\pi/2} \frac{1}{4} d\theta = \frac{\pi}{8}.$$

△

### 18.15.3 Coordonnées cylindriques

**Exemple 18.254.**

On veut calculer le volume de la région  $A$  définie par l'intersection entre la boule unité et le cylindre qui a pour base un disque de rayon  $1/2$  centré en  $(0, 1/2)$

$$A = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1\} \cap \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + (y - 1/2)^2 \leq 1/4\}.$$

On peut décrire  $A$  en coordonnées cylindriques

$$A = \left\{ (r, \theta, z) \in ]0, +\infty[ \times [-\pi, \pi[ \times \mathbb{R} \mid \right. \\ \left. -\pi/2 < \theta < \pi, 0 < r \leq \sin \theta, -\sqrt{1-r^2} \leq z \leq \sqrt{1-r^2} \right\}. \quad (18.852)$$

Le jacobien de ce changement de variables,  $J_{cyl}$ , est

$$J_{cyl}(r, \theta, z) = \begin{vmatrix} \cos \theta & \sin \theta & 0 \\ -r \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r, \quad (18.853)$$

qui est toujours positif. Le volume de  $A$  est donc

$$\int_{\mathbb{R}^3} \chi_A(x, y, z) dV = \int_{-\pi/2}^{\pi/2} \int_0^{\sin \theta} \int_{-\sqrt{1-r^2}}^{\sqrt{1-r^2}} r dz dr d\theta = \frac{2\pi}{8} + \frac{8}{9}.$$

△

**Exemple 18.255** (Volume d'un solide de révolution).

Soit  $g : [a, b] \rightarrow \mathbb{R}_+$  une fonction continue et positive. On dit que le solide  $A$  décrit par

$$A = \left\{ (x, y, z) \in \mathbb{R}^3 \mid z \in [a, b], \sqrt{x^2 + y^2} \leq g^2(z) \right\}$$



est un solide de révolution. Afin de calculer son volume, on peut décrire  $A$  en coordonnées cylindriques,

$$A = \{(r, \theta, z) \in ]0, +\infty[ \times [-\pi, \pi[ \times \mathbb{R} \mid a \leq z \leq b, 0 < r^2 \leq g^2(z)\}.$$

Le jacobien de ce changement de variables est  $J_{cyl} = r$ , comme dans l'exemple précédent. Le volume de  $A$  est donc

$$\int_{\mathbb{R}^3} \chi_A(x, y, z) dV = \int_a^b \int_{-\pi}^{\pi} \int_0^{g(z)} r dr d\theta dz = \int_a^b \pi g^2(z) dz.$$

Cette formule peut être utilisée pour tout solide de révolution.  $\triangle$

### 18.15.3.1 Coordonnées sphériques

Le calcul est un peu plus long :

$$\begin{aligned} J(\rho, \theta, \varphi) &= \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{vmatrix} \\ &= \begin{vmatrix} \sin \theta \cos \varphi & \rho \cos \theta \cos \varphi & -\rho \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & \rho \cos \theta \sin \varphi & -\rho \sin \theta \cos \varphi \\ \cos \theta & -\rho \sin \theta & 0 \end{vmatrix} \\ &= \rho^2 \sin \theta. \end{aligned} \quad (18.854)$$

Donc

$$dx dy dz = \rho^2 \sin(\theta) d\rho d\theta d\varphi. \quad (18.855)$$

#### Exemple 18.256.

On veut calculer le volume du cornet de glace  $A$

$$A = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in \mathbb{S}^2, \sqrt{x^2 + y^2} \leq z \leq \sqrt{1 - x^2 - y^2}\}.$$

On peut décrire  $A$  en coordonnées sphériques.

$$A = \{(\rho, \theta, \phi) \in ]0, +\infty[ \times [-\pi, \pi[ \times [0, \pi[ \mid 0 < \phi \leq \pi/4, 0 < \rho \leq 1\}.$$

Le jacobien de ce changement de variables  $J_{sph}$  est

$$J_{sph}(\rho, \theta, \phi) = \begin{vmatrix} \cos \theta \sin \phi & \sin \theta \sin \phi & \cos \phi \\ -\rho \sin \theta \sin \phi & \rho \cos \theta \sin \phi & 0 \\ \rho \cos \theta \cos \phi & \rho \sin \theta \cos \phi & -\rho \sin \phi \end{vmatrix} = \rho^2 \sin \phi, \quad (18.856)$$

Le volume de  $A$  est donc

$$\int_{\mathbb{R}^3} \chi_A(x, y, z) dV = \int_{-\pi}^{\pi} \int_0^{\pi/4} \int_0^1 \rho^2 \sin \phi d\rho d\phi d\theta = \frac{2\pi}{3} \left(1 - \frac{1}{\sqrt{2}}\right).$$

$\triangle$

### 18.15.4 Un autre système utile

Un changement de variables que l'on voit assez souvent est

$$\begin{cases} u = x + y \\ v = x - y. \end{cases} \quad (18.857a)$$

$$(18.857b)$$

Afin de calculer son jacobien, il faut d'abord exprimer  $x$  et  $y$  en fonctions de  $u$  et  $v$  :

$$\begin{cases} x = (u + v)/2 \\ y = (u - v)/2. \end{cases} \quad (18.858a)$$

$$(18.858b)$$

La matrice jacobienne est

$$\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}. \quad (18.859)$$

Le déterminant vaut  $-\frac{1}{2}$ . Nous avons donc

$$dxdy = \frac{1}{2}dudv. \quad (18.860)$$

Nous insistons sur le fait que c'est  $\frac{1}{2}$  et non  $-\frac{1}{2}$  qui intervient parce que la formule du changement de variable demande d'introduire la *valeur absolue* du jacobien.

**Exemple 18.257.**

Calculer l'intégrale de la fonction  $f(x, y) = x^2 - y^2$  sur le domaine représenté sur la figure 18.13.

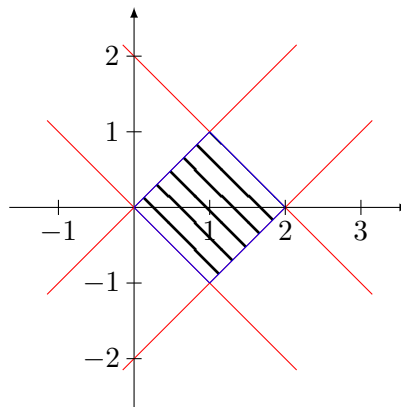


FIGURE 18.13: Un domaine qui s'écrit étonnamment bien avec un bon changement de coordonnées.

Les droites qui délimitent le domaine d'intégration sont

$$\begin{aligned} y &= -x + 2 \\ y &= x - 2 \\ y &= x \\ y &= -x \end{aligned} \quad (18.861)$$

Le domaine est donc donné par les équations

$$\begin{cases} y + x < 2 & (18.862a) \\ y - x > -2 & (18.862b) \\ y - x < 0 & (18.862c) \\ y + x > 0. & (18.862d) \end{cases}$$

En utilisant le changement de variables  $u = x + y$ ,  $v = x - y$  nous trouvons le domaine  $0 < u < 2$ ,  $0 < v < 2$ . En ce qui concerne la fonction,  $f(x, y) = (x + y)(x - y)$  et par conséquent

$$f(u, v) = uv. \quad (18.863)$$

L'intégrale à calculer est simplement

$$\int_0^2 \int_0^2 uv \, dudv = \int_0^2 u \, du \left[ \frac{v^2}{2} \right]_0^2 = 2 \int_0^2 u \, du = 4. \quad (18.864)$$

△

## 18.16 Aire d'une surface de révolution

### ⚠ Avertissement/question au lecteur !! 18.258

*Vu que nous n'avons pas encore de définition d'une aire dans  $\mathbb{R}^3$ , ce qui suit n'est en réalité pas très rigoureux.*

Soit  $\gamma$  une courbe dans le plan  $xy$ , paramétrée par

$$\gamma(u) = \begin{pmatrix} x(u) \\ y(u) \\ 0 \end{pmatrix} \quad (18.865)$$

avec  $u \in [a, b]$ . Nous supposons que la courbe est toujours positive, c'est-à-dire  $y(u) > 0$  pour tout  $u$ .

Nous voulons considérer la surface obtenue en effectuant une rotation de cette ligne autour de l'axe  $X$ . Chaque point de la courbe va parcourir un cercle de rayon  $y(u)$  dans le plan  $YX$  et centré en  $(x(u), 0, 0)$ . La surface est donc donnée par

$$\varphi(u, \theta) = \begin{pmatrix} x(u) \\ y(u) \cos \theta \\ y(u) \sin \theta \end{pmatrix} \quad (18.866)$$

avec  $(u, \theta) \in [a, b] \times [0, 2\pi]$ . Notez que la courbe de départ correspond à  $\theta = 0$ .

Les vecteurs tangents à la surface pour ce paramétrage sont

$$T_u = \frac{\partial \varphi}{\partial u} = \begin{pmatrix} x'(u) \\ y'(u) \cos \theta \\ y'(u) \sin \theta \end{pmatrix} \quad T_\theta = \frac{\partial \varphi}{\partial \theta} = \begin{pmatrix} 0 \\ -y(u) \sin \theta \\ y(u) \cos \theta \end{pmatrix}. \quad (18.867)$$

Le produit vectoriel de ces deux vecteurs vaut

$$\begin{aligned} T_u \times T_\theta &= \begin{vmatrix} e_x & e_y & e_z \\ x' & y' \cos \theta & y' \sin \theta \\ 0 & -y \sin \theta & y \cos \theta \end{vmatrix} \\ &= y'(u)y(u) e_x - x'(u)y(u) \cos \theta e_y + x'(u)y(u) \sin \theta e_z. \end{aligned} \quad (18.868)$$

En ce qui concerne la norme :

$$dS = \|T_u \times T_\theta\| = \sqrt{(y'y)^2 + (x'y)^2} = |y(u)|\sqrt{y'(u)^2 + x'(u)^2}. \quad (18.869)$$

Étant donné que nous avons supposé que  $y(u) > 0$ , nous pouvons supprimer les valeurs absolues, et l'aire de la surface de révolution devient :

$$\begin{aligned} \text{Aire}(S) &= \int_0^{2\pi} d\theta \int_a^b y(u) \sqrt{x'(u)^2 + y'(u)^2} du \\ &= 2\pi \int_a^b y(u) \sqrt{x'(u)^2 + y'(u)^2} du. \end{aligned} \quad (18.870)$$

### Exemple 18.259.

Calculons la surface du cône de révolution de rayon (à la base)  $R$  et de hauteur  $h$ . La courbe de départ est le segment droite qui part de  $(0, 0)$  et qui termine en  $(R, h)$  de la figure 18.14.

Ce segment peut être paramétré par

$$\gamma(u) = \begin{pmatrix} Ru \\ hu \\ 0 \end{pmatrix} \quad (18.871)$$

EXooZCLXooVmXQgY

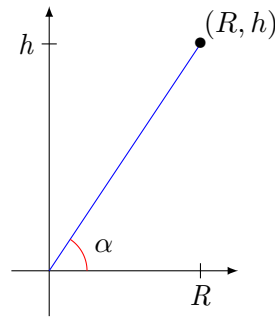


FIGURE 18.14: En faisant tourner cette droite autour de l'axe  $X$ , nous obtenons un cône.

avec  $u \in [0, 1]$ . Cela donne  $x(u) = Ru$ ,  $y(u) = hu$  et par conséquent

$$Aire = 2\pi \int_0^1 hu\sqrt{R^2 + h^2} = \pi h\sqrt{R^2 + h^2}. \quad (18.872)$$

Ce résultat peut aussi être exprimé en fonction de l'angle, grâce à la formule (18.580). En sachant que  $h = \sqrt{h^2 + R^2} \sin(\alpha)$ , nous trouvons

$$Aire = \pi(R^2 + h^2) \sin(\alpha). \quad (18.873)$$

△

### Exemple 18.260.

Calculons la surface latérale du tore obtenu par révolution du cercle de la figure 18.15.

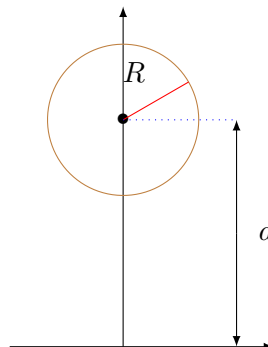


FIGURE 18.15: Si nous tournons ce cercle autour de l'axe  $X$ , nous obtenons un tore de rayon « externe »  $a$  et de rayon « interne »  $R$ .

Le chemin qui détermine le cercle de départ est

$$\gamma(u) = \begin{pmatrix} R \cos(u) \\ a + R \sin(u) \\ 0 \end{pmatrix}, \quad (18.874)$$

c'est-à-dire  $x(u) = R \cos(u)$ ,  $y(u) = a + R \sin(u)$  avec  $u \in [0, 2\pi]$ . Nous avons donc l'aire

$$\begin{aligned} Aire &= 2\pi \int_0^{2\pi} (a + R \sin(u)) R du \\ &= 2\pi R (2\pi a + R[-\cos(u)]_0^{2\pi}) \\ &= 4\pi^2 a R. \end{aligned} \quad (18.875)$$

△

## 18.17 Table de caractères du groupe diédral

Cette section vient de [126] ; nous avons comme but d'établir la table des caractères des représentations complexes du groupe diédral  $D_n$ . SecWMzheKf

### 18.17.1 Représentations de dimension un

Nous nous occupons des représentations de  $D_n$  sur  $\mathbb{C}$ . Les applications linéaires  $\mathbb{C} \rightarrow \mathbb{C}$  sont seulement les multiplications par des nombres complexes. Nous cherchons donc  $\psi: D_n \rightarrow \mathbb{C}^*$ .

Nous savons que  $D_n$  est généré<sup>89</sup> par  $s$  et  $r$ . Vu que  $s^2 = 1$ , nous avons

$$\psi(s)^2 = \psi(s^2) = \psi(1) = 1, \quad (18.876)$$

donc  $\psi(s) \in \{-1, 1\}$ . Nous savons aussi que  $sr sr = 1$ , donc

$$\psi(s)^2 \psi(r)^2 = 1, \quad (18.877)$$

ce qui donne  $\psi(r) \in \{-1, 1\}$ .

Nous avons donc quatre représentations de dimension un données par

|                | $\psi(r) = 1$ | $\psi(r) = -1$ |
|----------------|---------------|----------------|
| $\psi(s) = 1$  | $\rho^{++}$   | $\rho^{+-}$    |
| $\psi(s) = -1$ | $\rho^{-+}$   | $\rho^{--}$    |

Attention au fait que nous devons aussi avoir la relation  $\psi(r)^n = \psi(r^n) = 1$ . Donc  $\psi(r)$  doit être une racine  $n^e$  de l'unité. Nous allons donc devoir avoir un compte différent selon la parité de  $n$ . Nous en reparlerons à la fin, au moment de faire les comptes. En ce qui concerne les caractères correspondants,

|             | $r^k$    | $sr^k$       |
|-------------|----------|--------------|
| $\chi^{++}$ | 1        | 1            |
| $\chi^{+-}$ | $(-1)^k$ | $(-1)^k$     |
| $\chi^{-+}$ | 1        | -1           |
| $\chi^{--}$ | $(-1)^k$ | $(-1)^{k+1}$ |

Étant donné qu'ils sont tous différents, ce sont des représentations deux à deux non équivalentes, lemme 16.16.

### 18.17.2 Représentations de dimension deux

Nous cherchons maintenant les représentations  $\rho: D_n \rightarrow \text{End}(\mathbb{C}^2)$ . Ici nous supposons connue la liste des éléments de  $D_n$  donnée par le corolaire 18.175. Soit  $\omega = e^{2i\pi/n}$  et  $h \in \mathbb{Z}$  ; nous considérons la représentation  $\rho^{(h)}$  de  $D_n$  définie par

$$\rho^{(h)}(r^k) = \begin{pmatrix} \omega^{hk} & 0 \\ 0 & \omega^{-hk} \end{pmatrix} \quad (18.878a)$$

$$\rho^{(h)}(sr^k) = \begin{pmatrix} 0 & \omega^{-hk} \\ \omega^{hk} & 0 \end{pmatrix}. \quad (18.878b)$$

Cela donne bien  $\rho^{(h)}$  sur tous les éléments de  $D_n$  par la proposition 18.174. Nous pouvons restreindre le domaine de  $h$  en remarquant d'abord que  $\rho^{(h)} = \rho^{(h+n)}$ , et ensuite que les représentations  $\rho^{(h)}$  et  $\rho^{(-h)}$  sont équivalentes. Un opérateur d'entrelacement est donné par  $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ , et il est facile de vérifier que  $T\rho^{(h)}(x) = \rho^{(-h)}(x)T$  avec  $x = r^k$  puis avec  $x = sr^k$ .

Donc  $\rho^{(h)} \simeq \rho^{(-h)} \simeq \rho^{(n-h)}$  et nous pouvons restreindre notre étude à  $0 \leq h \leq \frac{n}{2}$ .

89. Voir proposition 18.174 et tout ce qui suit.

Nous allons séparer les cas  $n = 0$ ,  $h = n/2$  et les autres. En effet si nous notons par commodité  $a = \omega^h$ , alors un vecteur  $(x, y)$  est vecteur propre de  $\rho^{(h)}(s)$  et de  $\rho^{(h)}(r)$  si et seulement si il vérifie les systèmes d'équations

$$\begin{cases} ax = \lambda x \\ \frac{1}{a}y = \lambda y \end{cases} \quad (18.879a)$$

$$\begin{cases} \frac{1}{a}y = \lambda y \\ ax = \lambda y \end{cases} \quad (18.879b)$$

et

$$\begin{cases} \frac{1}{a}y = \mu x \\ ax = \mu y \end{cases} \quad (18.880a)$$

$$\begin{cases} \frac{1}{a}y = \mu x \\ ax = \mu y \end{cases} \quad (18.880b)$$

avec  $\lambda$  et  $\mu$  des nombres non nuls. Une représentation sera réductible si et seulement si ces deux systèmes acceptent une solution non nulle commune. Il est vite vu que si  $x \neq 0$  et  $y \neq 0$ , alors  $a^2 = 1$ , ce qui signifie  $h = 0$  ou  $h = n/2$ . Sinon, il n'y a pas de solutions, et la représentation associée est irréductible.

(1)  $h = 0$ . Nous avons

$$\rho^{(0)}(r^k) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \rho^{(0)}(sr^k) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (18.881)$$

donc le caractère de cette représentation est  $\chi^{(0)}(r^k) = 2$  et  $\chi^{(0)}(sr^k) = 0$ . Donc nous avons

$$\chi^{(0)} = \chi^{++} + \chi^{-+}. \quad (18.882)$$

Il y a maintenant (au moins) quatre façons de voir que la représentation  $\rho^{(0)}$  est réductible.

**Première méthode** Trouver un opérateur d'entrelacement. Pour cela nous calculons les matrices :

$$S(r^k) = (\rho^{++} \oplus \rho^{-+})(r^k) = \begin{pmatrix} \rho^{++}(r^k) & 0 \\ 0 & \rho^{-+}(r^k) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (18.883a)$$

$$S(sr^k) = (\rho^{++} \oplus \rho^{-+})(sr^k) = \begin{pmatrix} \rho^{++}(sr^k) & 0 \\ 0 & \rho^{-+}(sr^k) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (18.883b)$$

$$(18.883c)$$

Nous cherchons une matrice  $T$  telle que  $TS(r^k) = \rho^{(0)}(r^k)T$  et  $TS(sr^k) = \rho^{(0)}(sr^k)T$ . Étant donné que  $S(r^k) = \mathbb{1} = \rho^{(0)}(r^k)$ , la première contrainte n'en est pas une. Nous pouvons vérifier qu'avec  $T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ , nous avons bien

$$T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (18.884)$$

Donc ce  $T$  entrelace  $\rho^{++} \oplus \rho^{-+}$  avec  $\rho^{(0)}$  qui sont donc deux représentations équivalentes. Donc  $\rho^{(0)}$  est réductible et ça ne nous intéresse pas de la lister.

**Seconde méthode** Invoquer le théorème 16.24(1) et dire que les représentations sont équivalentes parce que les caractères sont égaux.

**Troisième méthode** Utiliser le théorème 16.24(2) et calculer  $\langle \chi^{(0)}, \chi^{(0)} \rangle$  :

$$\langle \chi^{(0)}, \chi^{(0)} \rangle = \frac{1}{|D_n|} \sum_{g \in D_n} |\chi^{(0)}(g)|^2 \quad (18.885a)$$

$$= \frac{1}{2n} (4 + 0 + 4(n-1)) \quad (18.885b)$$

$$= 2. \quad (18.885c)$$

Ici le 4 est pour le 1, le zéro est pour les termes  $sr^k$  et  $4(n-1)$  est pour les  $n-1$  termes  $r^k$ . Vu que le résultat n'est pas 1, la représentation  $\rho^{(0)}$  n'est pas irréductible.

**Quatrième méthode** Regarder les solutions des systèmes (18.879) et (18.880) dont nous avons parlé plus haut.

La première méthode a l'avantage d'être simple et ne demander aucune théorie particulière à part les définitions. La seconde méthode est la plus rapide, mais demande un théorème très puissant. La troisième utilise également un théorème assez avancé, mais a l'avantage sur les deux autres méthodes de ne pas avoir besoin de savoir à priori un candidat décomposition de  $\rho^{(0)}$ ; cette méthode est applicable même sans faire la remarque que  $\chi^{(0)} = \chi^{++} + \chi^{-+}$ .

Quoi qu'il en soit, nous ne listons pas  $\chi^{(0)}$  dans notre [table de caractères](#).

(2)  $h = n/2$ . Vu que  $\omega^{n/2} = e^{i\pi} = -1$ , nous avons

$$\rho^{(n/2)}(r^k) = \begin{pmatrix} (-1)^k & 0 \\ 0 & (-1)^k \end{pmatrix} \quad \rho^{(n/2)}(sr^k) = \begin{pmatrix} 0 & (-1)^k \\ (-1)^k & 0 \end{pmatrix}, \quad (18.886)$$

et donc

$$\chi^{(n/2)}(r^k) = 2(-1)^k \quad (18.887a)$$

$$\chi^{(n/2)}(sr^k) = 0. \quad (18.887b)$$

Il est vite vu que  $\chi^{(n/2)} = \chi^{+-} + \chi^{-+}$ . Ergo la représentation  $\rho^{(n/2)}$  n'est pas irréductible.

(3)  $0 < h < \frac{n}{2}$ . Dans ce cas nous avons  $\omega^h \neq \omega^{-h}$ , et en regardant les systèmes d'équations donnés plus haut, nous voyons que  $\rho^{(h)}(s)$  et  $\rho^{(h)}(r)$  n'ont pas de vecteurs propres communs. Donc ces représentations sont irréductibles.

Nous devons cependant encore vérifier si elles sont deux à deux non équivalentes. Supposons que pour  $h \neq h'$  nous ayons une matrice  $T \in \text{GL}(2, \mathbb{C})$  telle que  $T\rho^{(h)}(r)T^{-1} = \rho^{(h')}(r)$ . Cela impliquerait en particulier que les matrices  $\rho^{(h)}(r)$  et  $\rho^{(h')}(r)$  aient même valeurs propres. Nous aurions donc  $\{\omega^h, \omega^{-h}\} = \{\omega^{h'}, \omega^{-h'}\}$ . Mais cela est impossible avec  $0 < h < h' < \frac{n}{2}$ . Donc toutes ces représentations sont distinctes.

Le caractère de la représentation  $\rho^{(h)}$  est  $\chi^{(h)}(r^k) = \omega^{hk} + \omega^{-hk} = 2 \cos\left(\frac{2\pi hk}{n}\right)$ .

Nous ajoutons donc la ligne suivante à notre liste :

|              |                                        |        |
|--------------|----------------------------------------|--------|
|              | $r^k$                                  | $sr^k$ |
| $\chi^{(h)}$ | $2 \cos\left(\frac{2\pi hk}{n}\right)$ | 0      |

### 18.17.3 Le compte pour $n$ pair

Nous avons 4 représentations de dimension 1 puis  $\frac{n}{2} - 1$  représentations de dimension 2. En tout nous avons

$$\frac{n}{2} + 3 \quad (18.888)$$

représentations irréductibles modulo équivalence. Cela fait le compte en vertu des classes de conjugaisons listées en 18.10.13.5. Pour rappel, le nombre de représentations non équivalentes est égal au nombre de classes de conjugaison par le corolaire 16.27. Notons que c'est cela qui justifie le fait que nous ne devons pas chercher d'autres représentations. Nous sommes sûrs de les avoir toutes trouvées.

### 18.17.4 Le compte pour $n$ impair

Nous avons fait mention plus haut du fait que si  $\psi$  est une représentation de dimension 1, le nombre  $\psi(r)$  devait être une racine  $n^{\text{e}}$  de l'unité. Donc en dimension 1 nous avons seulement les représentations  $\rho^{++}$  et  $\rho^{-+}$ . Pour celles de dimension 2, nous en avons  $\frac{n-1}{2}$ . En tout nous avons donc

$$\frac{n+3}{2} \quad (18.889)$$

représentations irréductibles modulo équivalence. Cela fait le compte en vertu des classes de conjugaisons listées en 18.10.13.6.

# Chapitre 19

## Corps finis, racines de l'unité

### 19.1 Le groupe des racines de l'unité dans les nombres complexes

SecGJOLooWdMYVI

#### 19.1.1 Le groupe

DEFooDUWPooZaAByH

##### Définition 19.1.

Une **racine  $n^e$  de l'unité** dans un anneau est une racine du polynôme  $X^n - 1$ .

LEMooSXFBooYJmRTK

##### Lemme-Définition 19.2.

Soit

$$U_n = \{e^{2i\pi k/n} \text{ tel que } k = 0, \dots, n-1\} \quad (19.1)$$

EqIEAXooIpyFPe

- (1) L'ensemble  $U_n$  est un groupe pour la multiplication.
- (2) L'ensemble  $U_n$  est l'ensemble des racines  $n^e$  de l'unité dans  $\mathbb{C}$ .

*Démonstration.* Il est vite vérifié que tous les éléments de  $U_n$  sont des racines de l'unité parce que

$$(e^{2i\pi k/n})^n = e^{2i\pi k} = 1, \quad (19.2)$$

entre autres à cause du lemme 18.11.

Cela nous donne déjà  $n$  racines pour  $X^n - 1$  dans  $\mathbb{C}$ . Le théorème 6.110 nous indique qu'il ne peut pas y en avoir plus.  $\square$

#### 19.3.

En ce qui concerne les notations, dans  $U_n$ , le « U » signifie « unité ». Cela n'a à peu près rien à voir avec le « U » du groupe  $SU(n)$  ; dans ce dernier, le « U » est pour « unitaire ».

Un des intérêts du groupe des racines est qu'il permet de factoriser  $X^n - 1$ , comme nous le verrons via les polynômes cyclotomiques dans le lemme 19.20.

LEMooRQRYooWZAuSD

##### Lemme 19.4 ([622]).

Soient  $z_1, \dots, z_n \in \mathbb{C}$  tels que  $\sigma_k(z_1, \dots, z_n) \in \mathbb{Z}$  pour tout  $k = 1, \dots, n$ . Alors pour tout  $l \in \mathbb{N}$  et tout  $r \in \mathbb{N}^*$ , nous avons  $\sigma_r(z_1^l, \dots, z_n^l) \in \mathbb{Z}$ .

*Démonstration.* L'application  $(c_1, \dots, c_n) \mapsto \sigma_r(c_1^l, \dots, c_n^l)$  est un polynôme symétrique : il existe un polynôme symétrique  $P$  à coefficients dans  $\mathbb{Z}$  tel que

$$\sigma_r(z_1^l, \dots, z_n^l) = P(z_1, \dots, z_n). \quad (19.3)$$

Par le théorème de décomposition 6.180, il existe un polynôme  $Q$  à coefficients dans  $\mathbb{Z}$  tel que

$$P(z_1, \dots, z_n) = Q(\sigma_1(z_1, \dots, z_n), \dots, \sigma_n(z_1, \dots, z_n)). \quad (19.4)$$

Vu que  $Q$  est à coefficients entiers et que chacun des  $\sigma_i(z_1, \dots, z_n)$  est entier, le tout est entier.  $\square$



LEMooDJQFooOhVoPj

**Lemme 19.5** ([623]).

L'ensemble des polynômes unitaires de degré  $n$  dans  $\mathbb{Z}[X]$  dont les racines sont dans  $\overline{B(0,1)}$  (boule dans  $\mathbb{C}$ ) est fini.

*Démonstration.* Nous notons  $n$  le degré de  $P$ , et  $z_1, \dots, z_n$  les racines de  $P$ . Les relations coefficients-racines 6.183 dit que

$$\sigma_k(z_1, \dots, z_n) = (-1)^k a_{n-k} \quad (19.5)$$

où  $\sigma_k$  est le polynôme symétrique élémentaire défini en 6.179. Vu que  $P \in \mathbb{Z}[X]$ , nous avons  $\sigma_k(z_1, \dots, z_n) \in \mathbb{Z}$ , et pour chaque  $k$  nous avons

$$|\sigma_k(z_1, \dots, z_n)| = \left| \sum_{s \in F_k} \prod_{i=1}^k g_{s(i)} \right| \leq \sum_{s \in F_k} \prod_{i=1}^k |g_{s(i)}| \leq \sum_{s \in F_k} 1 = \text{Card}(F_k). \quad (19.6)$$

Nous avons montré que  $|a_{n-k}| \leq \text{Card}(F_k)$ . Il n'y a qu'un nombre fini de possibilités.  $\square$

PROPooINGJooPkByVP

**Proposition 19.6** (Théorème de Kronecker[623, 624, 622]).

Soit un polynôme unitaire de  $\mathbb{Z}[X]$  dont toutes les racines sont dans  $\overline{B(0,1)}$ . Si  $P(0) \neq 0$  alors toutes les racines de  $P$  sont des racines de l'unité.

*Démonstration.* Nous notons  $\Omega_n$  l'ensemble des polynômes unitaires dans  $\mathbb{Z}[X]$  de degré  $n$  dont les racines sont dans  $\overline{B(0,1)}$ . Ce ensemble est fini par le lemme 19.5.

Soit un polynôme unitaire de  $\mathbb{Z}[X]$  dont toutes les racines sont dans  $\overline{B(0,1)}$  et tel que  $P(0) \neq 0$ . Nous devons prouver que toutes les racines de  $P$  sont des racines de l'unité. Soient  $z_1, \dots, z_n$  les racines de  $P$ .

(1) **Les polynômes  $P_k$**  Pour chaque  $k \in \mathbb{N}^*$ , nous posons

$$P_k(X) = \prod_{i=1}^n (X - z_i^k). \quad (19.7)$$

C'est un polynôme symétrique unitaire de degré  $n$  dont les racines sont les  $z_i^k$  qui vérifient  $|z_i^k| \leq 1$ .

(2)  **$P_k \in \mathbb{Z}[X]$**  Nous montrons que les polynômes  $P_k$  sont à coefficients entiers. Les relations coefficients racines montrent que les coefficients de  $P_k$  sont essentiellement les nombres  $\sigma_r(z_1^k, \dots, z_n^k)$ , qui sont entiers par le lemme 19.4.

Nous avons prouvé que  $P_k \in \Omega_n$  pour tout  $k$ .

Notons que  $\Omega_n$  étant fini et chaque élément de  $\Omega_n$  ayant au maximum  $n$  racines distinctes, l'ensemble des racines d'éléments de  $\Omega_n$  est fini. Par ailleurs, pour chaque  $i$ , tous les  $z_i^k$  sont racines d'éléments de  $\Omega_n$  (en l'occurrence de  $P_k$ ). Par une autre manière de le dire est que  $\{z_i^k\}_{k \in \mathbb{N}}$  est un ensemble fini<sup>1</sup>. Il existe donc  $k \neq l$  tels que  $z_i^k = z_i^l$ . Nous avons alors  $z_i^{k-l} = 1$ , et  $z_i$  est une racine de l'unité.  $\square$

LemWHQGooXyeJiw

**Lemme 19.7.**

L'ensemble  $U_n$  est un groupe cyclique<sup>2</sup> d'ordre  $n$  généré par  $\xi = e^{2i\pi/n}$ .

*Démonstration.* Il y a trois propriétés à vérifier pour que ce soit un groupe.

- (1) **Neutre** Le nombre 1 est une racine de l'unité.
- (2) **Inverse** Si  $\omega \in U_n$  alors  $\omega^n = 1$  et donc  $\omega \omega^{n-1} = 1$ , ce qui signifie que  $\omega^{n-1}$  est un inverse de  $\omega$ . Il reste à voir que  $\omega^{n-1} \in U_n$ . En effet  $(\omega^{n-1})^n = (\omega^n)^{n-1} = 1^{n-1} = 1$ .
- (3) **Associativité** Cas particulier de l'associativité dans  $\mathbb{C}$ .

1. En tant que partie d'un ensemble fini. On pourrait invoquer le principe des tiroirs, mais la proposition 1.70 fait aussi bien l'affaire.

2. Définition 1.428.

Le fait que ce soit un groupe cyclique contenant  $n$  éléments fait partie de la définition.  $\square$

Le lemme suivant donne les autres générateurs.

LemcFTNMa

**Lemme 19.8.**

Le nombre  $\xi^a$  est un générateur de  $U_n$  si et seulement si  $\text{pgcd}(a, n) = 1$ .

*Démonstration.* Si  $\text{pgcd}(a, n) = 1$  alors le théorème de Bézout 1.349 nous fournit des entiers  $u$  et  $v$  tels que  $ua + vn = 1$ . Alors nous avons

$$e^{2i\pi/n} = e^{2(ua+vn)i\pi/n} = (e^{2ai\pi/n})^u, \quad (19.8)$$

ce qui signifie que  $\xi$  est dans le groupe engendré par  $\xi^a$ , et par conséquent tout  $U_n$  est engendré.

Pour l'implication inverse, nous utilisons le théorème de Bézout dans le sens inverse. Soit  $\xi^a$  un générateur de  $U_n$ . Alors il existe  $u$  tel que  $(\xi^a)^u = \xi$ , donc  $\xi^{au-1} = 1$ , c'est-à-dire qu'il existe  $v$  tel que  $au - 1 = vn$ . Cette dernière égalité implique que  $\text{pgcd}(a, n) = 1$ .  $\square$

**Exemple 19.9.**

Une conséquence tout à fait extraordinaire de ce lemme est que le nombre 7 est générateur de  $\mathbb{Z}/12\mathbb{Z}$  (parce que  $\text{pgcd}(7, 12) = 1$ ). Or en solfège, une quinte fait 7 demi-tons, et une gamme en fait 12. Le cycle des quintes est donc générateur de la gamme chromatique[625]. Ce fait est connu des musiciens<sup>3</sup> depuis des siècles.  $\triangle$

PROPooIOQEooGmCJm

**Proposition 19.10** (Intersection par deux).

Les ensembles  $U_\alpha$  et  $U_\beta$  ont une intersection réduite à  $\{1\}$  si et seulement si  $\alpha$  et  $\beta$  sont premiers entre eux.

*Démonstration.* Nous rappelons qu'une racine  $\alpha^e$  de l'unité peut s'écrire sous la forme  $e^{2i\pi k/\alpha}$  avec  $0 \leq k < \alpha$ .

- (1) **Sens direct** Par contraposée, nous supposons que  $\alpha$  et  $\beta$  ne sont pas premiers entre eux, et nous notons  $d$  leur pgcd. Nous nommons  $\alpha = d\alpha'$  et  $\beta = d\beta'$ . Pour trouver une intersection entre  $U_\alpha$  et  $U_\beta$  nous devons trouver une valeur de  $0 < k < \alpha$  telle que

$$(e^{2i\pi k/\alpha})^\beta = e^{2i\pi k\beta/\alpha} = 1, \quad (19.9)$$

c'est-à-dire une valeur de  $k$  telle que  $k\beta/\alpha$  soit un entier. Mais  $k\beta/\alpha = k\beta'/\alpha'$  et par conséquent prendre  $k = \alpha'$  fonctionne. Surtout que par hypothèse  $d > 1$  et donc  $k = \alpha' < \alpha$ .

- (2) **Sens réciproque** Supposons maintenant que  $\alpha$  et  $\beta$  soient premiers entre eux. Soit  $z \in U_\alpha \cap U_\beta$ . Le fait que  $z$  soit une racine  $\alpha^e$  de l'unité implique qu'il existe un  $k < \alpha$  tel que  $z = e^{2i\pi k/\alpha}$ . Mais si  $z$  est également une racine  $\beta^e$  de l'unité, alors  $z^\beta = 1$ , c'est-à-dire que  $k\beta/\alpha$  doit être un entier, soit  $l$  cet entier. Nous avons

$$k\beta = l\alpha. \quad (19.10)$$

Si  $k > 0$ , comme le nombre  $\alpha$  divise  $k\beta$ , cela conduirait via le lemme de Gauss 1.361 à dire que  $\alpha$  divise  $k$ . Mais  $\alpha$  ne peut pas diviser  $k$  parce que nous avons supposé que  $k$  était strictement plus petit que  $\alpha$ . Donc  $k = 0$  et  $z = 1$ .  $\square$

PropFDDHooEyYxBC

**Proposition 19.11** (Intersection : le cas général[1]).

Soient des entiers positifs  $\alpha_1, \dots, \alpha_p$ . Nous avons

$$\bigcap_{i=1}^p U_{\alpha_i} = \{1\} \quad (19.11)$$

si et seulement si  $\text{pgcd}(\alpha_1, \dots, \alpha_p) = 1$  (c'est-à-dire que les  $\alpha_i$  sont premiers dans leur ensemble).

3. Même ceux qui ignorent le théorème de Bézout.

*Démonstration.* Nous décomposons les  $\alpha_i$  en facteurs premiers<sup>4</sup> de la façon suivante :  $\alpha_i = \prod_{k \in \mathbb{N}} p_k^{\alpha_i^{(k)}}$  où les  $p_k$  sont les nombres premiers.

- (1) **Caractérisation par une décomposition en facteurs premiers** Les éléments  $z$  différents de 1 dans  $U_{\alpha_1}$  s'écrivent sous la forme

$$z = e^{2i\pi k/\alpha_1} \quad (19.12)$$

avec  $0 < k < \alpha_1$ .

Pour tout  $i \neq 1$ , le fait que  $z \in U_{\alpha_i} \cap U_{\alpha_1}$  se traduit par le fait que  $(e^{2i\pi k/\alpha_1})^{\alpha_i} = 1$ , c'est-à-dire que  $\alpha_i k/\alpha_1$  est entier, donc que  $\alpha_1$  divise  $k\alpha_i$ . Par conséquent il existera un élément différent de 1 dans l'intersection des  $U_{\alpha_i}$  si et seulement si il existe un entier  $k$  strictement compris entre 0 et  $\alpha_1$  pour lequel  $\alpha_1$  divise tous les  $k\alpha_i$ .

Un entier  $0 < k < \alpha_1$  convient si et seulement si pour tout  $l$ , la puissance de  $p_l$  dans la décomposition de  $k$  est au moins égale à

$$\alpha_1^{(l)} - \alpha_i^{(l)} \quad (19.13)$$

pour tout  $l$ .

- (2) **Sens direct** L'hypothèse  $\text{pgcd}(\alpha_1, \dots, \alpha_p) \neq 1$  implique qu'il existe un  $l$  pour lequel tous les  $\alpha_i^{(l)}$  sont non nuls. Nous construisons le  $k$  voulu en prenant pour tout  $p_i$  la même puissance que celle dans  $\alpha_1$ , sauf pour  $p_l$  pour lequel nous prenons la puissance  $\alpha_1^{(l)} - \min_i \{\alpha_i^{(l)}\}$ . Le minimum en question est strictement positif, ce qui donne un  $k$  strictement inférieur à  $\alpha_1$ .
- (3) **Sens réciproque** Si  $\text{pgcd}(\alpha_1, \dots, \alpha_p) = 1$  alors pour tout  $l$ , il existe un  $i$  tel que  $\alpha_i^{(l)} = 0$ . Donc pour tout  $l$ , la puissance de  $p_l$  dans la décomposition de  $k$  est au moins  $\alpha_1^{(l)}$ . Cela implique que  $k \geq \alpha_1$ , ce qui est impossible.

□

### 19.1.2 Fonction indicatrice d'Euler

Voir le thème 11.

DefLYGTTooFPOYGZ

#### Définition 19.12.

Les générateurs de  $U_n$  sont les **racines primitives**<sup>5</sup> de l'unité dans  $\mathbb{C}$ . Nous nommons  $\Delta_n$  leur ensemble :

$$\Delta_n = \{e^{2i\pi k/n} \text{ tel que } 1 \leq k \leq n \text{ et } \text{pgcd}(k, n) = 1\}. \quad (19.14)$$

Nous avons par exemple

- Pour  $n = 1$ ,  $\text{pgcd}(k, 1) = 1$  et  $1 \leq k \leq 1$  uniquement pour  $k = 1$ .
- Pour  $n = 2$ ,  $\text{pgcd}(k, 2) = 1$  et  $1 \leq k \leq 2$  uniquement pour  $k = 1$ .
- Pour  $n = 3$ ,  $\text{pgcd}(k, 3) = 1$  et  $1 \leq k \leq 3$  uniquement pour  $k = 1, k = 2$ .

$$\Delta_1 = \{1\} \quad (19.15a)$$

$$\Delta_2 = \{e^{i\pi}\} \quad (19.15b)$$

$$\Delta_4 = \{e^{i\pi/2}, e^{i3\pi/2}\}. \quad (19.15c)$$

Notons que  $1 \in \Delta_d$  seulement avec  $d = 1$ .

4. Théorème 1.365.

5. parce qu'en prenant les puissances successives de l'une d'entre elles, nous retrouvons toutes les racines de l'unité, voir aussi la définition 19.56.

PROPOWILSooFWLnt

**Proposition 19.13.**En posant<sup>6</sup>

$$U_n = \{e^{2i\pi k/n} \text{ tel que } k = 0, \dots, n-1\}, \quad (19.16)$$

et<sup>7</sup>

$$\Delta_d = \{e^{2i\pi k/d} \text{ tel que } 1 \leq k \leq d \text{ et } \text{pgcd}(k, d) = 1\}, \quad (19.17)$$

nous avons

$$U_n = \bigcup_{d|n} \Delta_d \quad (19.18) \quad \text{EqZuLyL}$$

et l'union est disjointe.

*Démonstration.* En plein de petites parties.

- (1) **Union disjointe** Soient  $d_1, d_2$  tels que  $x \in \Delta_{d_1} \cap \Delta_{d_2}$ . Il existe  $k, l$  tels que  $e^{2i\pi k/d_1} = e^{2i\pi l/d_2}$ . Le corolaire 18.25 dit alors qu'il existe  $s \in \mathbb{N}$  tel que

$$\frac{k}{d_1} = \frac{l}{d_2} + s = \frac{l + sd_2}{d_2}. \quad (19.19)$$

Vu que  $k/d_2$  est une fraction irréductible, nous avons  $d_2 \leq d_1$  par la proposition 3.30. Nous faisons de même en permutant les rôles de  $d_1$  et  $d_2$  : il existe  $t \in \mathbb{N}$  tel que

$$\frac{l}{d_1} = \frac{k + d_1 t}{d_1}, \quad (19.20)$$

de telle sorte que  $d_1 \leq d_2$ . Nous en déduisons que  $d_1 = d_2$ . Donc l'union à droite dans (19.18) est disjointe.

- (2) **Première inclusion** Soit  $x \in U_n$ . Il existe  $0 \leq k \leq n-1$  tel que  $x = e^{2i\pi k/n}$ . Si  $x = 1$ , alors  $x \in \Delta_1$  ; nous supposons que ce n'est pas le cas, c'est-à-dire que  $k \neq 0$ . Nous écrivons la fraction  $k/n$  sous forme irréductible :

$$\frac{k}{n} = \frac{k'}{d} \quad (19.21)$$

avec  $d = \text{pgcd}(k, n)$ , et  $\text{pgcd}(k', d) = 1$ . Donc  $x \in \Delta_d$ .

- (3) **Seconde inclusion** Nous supposons que  $x \in \Delta_d$  pour un certain diviseur  $d$  de  $n$ . Il existe donc  $1 \leq k \leq d$  tel que  $\text{pgcd}(k, d) = 1$  et  $x = e^{2i\pi k/d}$ . Si  $d = 1$  alors  $x = 1$  et  $x \in U_n$ . Nous excluons ce cas. Vu que  $d \mid n$ , nous considérons  $s \in \mathbb{N}$  tel que  $n = sd$ . Nous avons alors

$$\frac{k}{d} = \frac{ks}{sd} = \frac{ks}{n}. \quad (19.22)$$

Multiplions par  $s$  les inéquations  $1 \leq k \leq d$  :

$$s \leq ks \leq sd = n. \quad (19.23)$$

Le cas  $ks = n$  est exclu parce que cela demanderait  $kn = ksd = nd$ , et donc  $k = d$ , ce que nous avons exclu parce que  $k = d$  est uniquement possible pour  $d = 1$  à cause de la condition  $\text{pgcd}(k, d) = 1$ .

□

PROPOoFKCHooWdFieM

**Proposition 19.14.**

Nous avons

$$\varphi(n) = \text{Card}(\Delta_n) \quad (19.24) \quad \text{EqEulerGqPsvi}$$

 $\varphi$  est l'indicatrice d'Euler<sup>8</sup>.

---

6. Voir la définition 19.2.

7. Définition 19.12. On pourrait croire que la condition  $\text{pgcd}(k, d) = 1$  exclu  $k = d$ . On aurait alors envie d'écrire  $1 \leq k < d$ . Mais pour  $d = 1$ , ça fait une différence parce que tel qu'écrit  $\Delta_1 = \{1\}$ .

8. Définition 5.29.

*Démonstration.* Nous prouvons que l'application

$$\begin{aligned} f: \{k \in \mathbb{N}^* \text{ tel que } 1 \leq k \leq n, \text{pgcd}(k, n) = 1\} &\rightarrow \{e^{2i\pi k/n} \text{ tel que } 1 \leq k \leq n, \text{pgcd}(k, n) = 1\} \\ k &\mapsto e^{2i\pi k/n} \end{aligned} \quad (19.25)$$

est une bijection.

Le fait que ce soit surjectif est immédiat. Pour l'injectivité, nous considérons  $1 \leq k \leq n$  et  $1 \leq l \leq n$  tels que  $f(k) = f(l)$ . Nous avons donc  $e^{2i\pi k/n} = e^{2i\pi l/n}$ . Le corolaire 18.25 nous dit alors qu'il existe  $s \in \mathbb{Z}$  tel que

$$2\pi k/n = 2\pi l/n + 2\pi s. \quad (19.26)$$

En simplifiant par  $2\pi$  et en multipliant par  $n$ , nous trouvons  $k = l + sn$ . Étant donné que  $k$  et  $l$  sont tout deux entre 1 et  $n$ , nous avons obligatoirement  $s = 0$  et donc  $k = l$ .

Donc  $f$  est bijective et  $\text{Card}(\Delta_n) = \varphi(n)$ . □

### 19.1.3 Décomposition en éléments simples

SUBSEMOOBNXdoBIDHUpdV

**Lemme 19.15** ([626, 1]).

Soient des nombres complexes distincts  $a_1, \dots, a_N$ . Alors pour tout  $z \notin \{a_i\}_{i=1, \dots, N}$ ,

$$\prod_{i=1}^N \frac{1}{z - a_i} = \sum_{i=1}^N \frac{\lambda_i}{z - a_i} \quad (19.27)$$

où

$$\lambda_i = \prod_{j \neq i} \frac{1}{a_i - a_j}. \quad (19.28)$$

*Démonstration.* Nous posons  $P(z) = \prod_{k=1}^N (z - a_k)$ , et nous calculons :

$$\sum_{i=1}^N \frac{\lambda_i}{z - a_i} = \sum_{i=1}^N \frac{\lambda_i \prod_{k \neq i} (z - a_k)}{\prod_{k=1}^N (z - a_k)} \quad (19.29a)$$

$$= \frac{1}{P(z)} \sum_{i=1}^N \lambda_i \prod_{k \neq i} (z - a_k) \quad (19.29b)$$

Il s'agit maintenant de prouver que la somme de gauche vaut toujours 1. Nous posons

$$S(z) = \sum_{i=1}^N \frac{\prod_{k \neq i} (z - a_k)}{\prod_{k \neq i} (a_i - a_k)}. \quad (19.30)$$

Calculons  $S(a_l)$  :

$$S(a_l) = \sum_{i=1}^N \frac{\prod_{k \neq i} (a_l - a_k)}{\prod_{k \neq i} (a_i - a_k)}. \quad (19.31)$$

Pour les termes  $i \neq l$ , le numérateur est nul, car il contient le facteur  $a_l - a_l = 0$ . Donc la somme se réduit au seul terme  $i = l$  :

$$S(a_l) = \frac{\prod_{k \neq l} (a_l - a_k)}{\prod_{k \neq l} (a_l - a_k)} = 1. \quad (19.32)$$

Le polynôme  $S - 1$  est donc un polynôme de degré  $N - 1$  qui possède  $N$  racines distinctes. Le théorème 6.110 implique que  $S - 1 = 0$  et donc que  $S = 1$  comme nous le voulions. □

Il est possible de décomposer une fraction rationnelle en fractions dites « simples ». Si  $|z| < 1$  nous avons par exemple la décomposition

$$\frac{1}{1 - z^r} = \sum_{\omega \in U_r} \frac{A_\omega}{\omega - z} \quad (19.33)$$

EqDWYBooJIMPAz

où  $U_r$  est le groupe des racines  $r^e$  de l'unité défini en (19.1). Les nombres  $A_\omega$  peuvent alors être déterminés en effectuant la somme. Le dénominateur commun sera  $1 - z^r$  tandis que les  $A_\omega$  sont déterminés en égalant le numérateur à 1.

**Exemple 19.16.**

Pour décomposer la fraction  $\frac{1}{1-x^2}$  nous savons que les racines du dénominateur sont  $\pm 1$ . Donc nous écrivons

$$\frac{1}{1-x^2} = \frac{A}{1-x} + \frac{B}{1+x}. \quad (19.34)$$

Nous trouvons les valeurs de  $A$  et  $B$  en effectuant la somme :

$$\frac{A(1+x) + B(1-x)}{1-x^2} = \frac{A+B + (A-B)x}{1-x^2}. \quad (19.35)$$

Les coefficients  $A$  et  $B$  doivent donc vérifier  $A+B=1$  et  $A-B=0$ . Au final,

$$\frac{1}{1-x^2} = \frac{1}{2(1-x)} + \frac{1}{2(1+x)}. \quad (19.36)$$

△

## 19.2 Chiffrement RSA

SecEVaFYi

Ce passage sur RSA provient en bonne partie de la page de Wikipédia[627].

Alice veut envoyer un message à Bob. L'idée est que Bob va donner à Alice une clef publique qui va permettre de chiffrer le message, tandis que Bob va garder pour lui une clef privée qui permet de déchiffrer.

### 19.2.1 Mise en place par Bob

Bob se crée une paire de clef publique, clef privée de la façon suivante.

- (1) Bob choisit deux nombres premiers distincts  $p, q$ .
- (2) Il calcule  $n = pq$ .
- (3) Par le corolaire 5.43, l'indicatrice d'Euler  $\varphi(n) = (p-1)(q-1)$  est facile à calculer pour Bob.
- (4) Bob choisit  $e \in \mathbb{N}$  premier avec  $\varphi(n)$ , puis  $d$  tel que  $ed \in [1]_{\varphi(n)}$ .

Maintenant la paire est : clef publique  $(n, e)$  et clef privée  $(n, d)$ <sup>9</sup>.

Bob envoie la paire  $(n, e)$  à Alice.

**Remarque 19.17.**

Ici nous ne supposons pas que la communication soit sûre. Une tierce personne peut intercepter le message. D'ailleurs en principe, les gens publient leur clef publique sur leurs sites, voire sur des sites dédiés. Le problème de l'identification reste à résoudre à l'ancienne.

### 19.2.2 Chiffrement

Nous chiffons en utilisant la clef publique  $(n, e)$ . D'abord Alice se débrouille pour transformer son message en un nombre plus petit que  $n$ . Soit  $M$  ce message. Alice code  $M$  en

$$C = M^e \mod n. \quad (19.37)$$

Tout le truc est que nous allons voir que l'application  $x \mapsto x^e$  est une bijection de  $\mathbb{F}_n$ , et que l'inverse est facile à calculer par Bob, et difficile pour les autres. Alice envoie  $C$  à Bob. Encore une fois, nous ne supposons pas que cette communication soit privée. Le nombre  $C$  peut être intercepté.

9. Le fait que  $e$  soit public et  $d$  soit privé est une convention.  $e$  comme *encryption* et  $d$  comme *decryption*.

### 19.2.3 Déchiffrement

Nous allons montrer que  $M = C^d \pmod n$ , et donc que Bob, connaissant  $(n, d)$ , peut déchiffrer. D'abord

$$C^d = (M^e)^d = M^{ed}, \quad (19.38)$$

mais nous savons qu'il existe  $k$  tel que

$$ed = 1 + k\varphi(n) = 1 + k(p-1)(q-1). \quad (19.39)$$

L'étape astucieuse est de remarquer que

$$M^{1+k(p-1)(q-1)} \in [M]_p \cap [M]_q. \quad (19.40)$$

Pour montrer cela nous utilisons le petit théorème de Fermat 6.15(4).

— Si  $M$  est premier avec  $p$ , alors  $M^{p-1} \in [1]_p$ .

— Si  $M$  n'est pas premier avec  $p$ , alors  $M$  est multiple de  $p$  et on sait que  $M^{p-1} \in [0]_p = [M]_p$ .

Dans les deux cas nous avons (19.40). Le nombre  $M^{1+k\varphi(n)} - M$  est donc à la fois multiple de  $p$  et de  $q$ .

Le lemme chinois 6.28 nous dit immédiatement<sup>10</sup> qu'alors

$$M^{1+k\varphi(n)} - M \quad (19.41)$$

est un multiple de  $pq = n$ , c'est-à-dire que

$$C^d = M^{ed} \in [M]_n. \quad (19.42)$$

Si on ne croit pas au lemme chinois, on peut utiliser le lemme de Gauss. Posons

$$M^{1+k\varphi(n)} - M = ap = bq. \quad (19.43)$$

Dans ce cas  $p$  divise  $bq$ , mais  $q$  est premier avec  $p$ , donc le lemme de Gauss 3.60 nous enseigne<sup>11</sup> que  $p$  divise  $b$ .

### 19.2.4 Une imprudence à ne pas commettre

Nous avons pris deux cas selon que  $M$  soit ou non premier avec  $p$ . Une question qui se pose est la suivante : est-ce que c'est une bonne idée d'envoyer un message qui ne soit pas premier avec  $p$  ?

Si nous savons que  $M$  n'est pas premier avec  $p$ , alors nous avons  $M^e = l^e p^e$  et  $n = pq$  qui sont publics. Donc un calcul de PGCD permettrait de trouver  $p$ .

Il faut cependant savoir que |

- La probabilité que ça arrive est infime : comme  $M$  est entre 0 et  $n = pq$ , les multiples de  $p$  possibles sont  $p, 2p, \dots, pq$ . Il y a donc une chance sur  $p$  que cela arrive. Typiquement avec des  $p$  de l'ordre de  $10^{120}$ , on peut utiliser RSA chaque milliseconde sur chaque atome de l'univers depuis le début des temps que ça ne se serait presque certainement pas encore produit.
- De toutes façons, Alice ne sait pas vérifier si son message est premier avec  $p$ , parce qu'elle ne connaît pas  $p$ .
- En conclusion, la partie de la preuve qui montre que  $M^{1+\varphi(n)} \in [M]_p \cap [M]_q$  dans le cas  $M$  non premier avec  $p$  est, à toutes fins pratiques, inutile parce que ce cas de figure ne se présentera jamais dans toutes l'histoire de l'univers, même pas avec une civilisation intelligente autour de chaque étoile.

ProbGAYFooZATuYy

### ⚠ Avertissement/question à la lectrice !! 19.18

*Est-ce que ces trois points sont corrects ?*

10. C'est ici qu'il est important que  $p$  ne soit pas égal à  $q$ . Si  $p = q$ , alors le lemme chinois ne fonctionne pas.

11. Ici aussi, si  $p = q$ , ça ne marche pas.

### 19.2.5 Problèmes calculatoires

Pour implémenter RSA, il faut pouvoir faire (au moins) trois choses :

- (1) Trouver de grands nombres premiers.
- (2) Trouver des couples de Bézout.
- (3) Calculer  $M^e$  lorsque  $e$  est très grand.

En ce qui concerne le problème de trouver des nombres premiers, c'est compliqué, mais il faut savoir qu'il y en a plein. À 120 chiffres, il y a environ autant de nombres premiers que d'atomes dans  $10^{20}$  fois l'univers connu. Cela rend impossible toute tentative de factoriser un grand nombre en essayant toutes les possibilités. Même pas en science-fiction.

Trouver des nombres  $u$  et  $v$  tels que  $Au + Bv = \text{pgcd}(A, B)$  est un problème expliqué en 3.2.1.

En ce qui concerne le calcul de  $M^e$  lorsque  $e$  est grand, il n'est évidemment pas pensable de faire  $M \cdot M \cdot \dots M$  avec  $e$  facteurs. Un truc pour calculer en moins d'étapes est l'**exponentiation rapide**. Si  $e = 2k$  est pair, nous calculons

$$M^e = (M^k)^2; \quad (19.44)$$

si  $e = 2k + 1$  alors nous calculons

$$M^e = M(M^k)^2. \quad (19.45)$$

Le calcul prend alors seulement environ  $\log_2(e)$  étapes. Pour donner une idée,

$$\log_2(10^{120}) \simeq 400. \quad (19.46)$$

Très raisonnable, mais un ordinateur reste indispensable.

### 19.2.6 La solidité de RSA

La solidité de la méthode repose sur deux conjectures (non démontrées!!) :

- Pour déchiffrer il *faut* connaître  $p$  et  $q$ .
- La difficulté de trouver  $p$  et  $q$  en partant de  $n = pq$  est exponentielle en  $n$ .

Dans la méthode de déchiffrement proposée ici,  $p$  et  $q$  sont utilisés pour calculer  $d$  qui est solution de  $ed = [1]_{\varphi(n)}$ . La seule formule connue pour calculer  $\varphi(n)$  est  $\varphi(n) = (p-1)(q-1)$ . Si on trouve plus simple, alors RSA peut être craqué.

### 19.2.7 Note non mathématique pour doucher l'enthousiasme

Il est souvent dit[628] que différents systèmes de chiffrement peuvent aider à avoir des discussions « discrètes » dans les régimes totalitaires. La technologie au service de la démocratie, voilà qui enthousiasme la jeunesse<sup>12</sup>. La réalité est qu'il est souvent possible de craquer un système de chiffrement arbitrairement complexe, même sans connaître le petit théorème de Fermat ... tout dépend du contexte.

## 19.3 Polynômes cyclotomiques

### 19.3.1 Définitions et propriétés

#### Définition 19.19.

Le *polynôme cyclotomique* d'indice  $n$  est le polynôme

$$\phi_n(X) = \prod_{z \in \Delta_n} (X - z) \quad (19.47)$$

12. Cela dit, le navigateur Tor[629], qui est un pur produit de RSA, permet effectivement d'accéder en France aux sites bloqués pour apologie du terrorisme (mars 2015).



où  $\Delta_n$  est l'ensemble des racines primitives de l'unité de la définition 19.12 :

$$\Delta_n = \{e^{2i\pi k/n} \text{ tel que } 1 \leq k \leq n \text{ et } \text{pgcd}(k, n) = 1\}, \quad (19.48)$$

voir la définition 19.12.

Le polynôme  $\phi_n$  est un polynôme unitaire de degré  $\varphi(n)$  où  $\varphi$  est l'indicatrice d'Euler <sup>13</sup>. Nous avons par exemple

$$\Delta_1 = \{1\} \quad (19.49a)$$

$$\Delta_2 = \{-1\} \quad (19.49b)$$

$$\Delta_3 = \{e^{2i\pi/3}, e^{4i\pi/3}\} \quad (19.49c)$$

et les premiers polynômes cyclotomiques sont donnés par

$$\phi_1(X) = X - 1 \quad (19.50a)$$

$$\phi_2(X) = X + 1 \quad (19.50b)$$

$$\phi_3(X) = X^2 + X + 1. \quad (19.50c)$$

Pour le dernier nous avons utilisé le fait que  $e^{6i\pi/3} = 1$  et  $e^{2i\pi/3} + e^{4i\pi/3} = -1$ .

LemKYGBooAwpOHD

#### Lemme 19.20.

Le polynôme  $X^n - 1$  se factorise des diverses manières suivantes :

$$X^n - 1 = \prod_{z \in U_n} (X - z) = \prod_{d|n} \prod_{z \in \Delta_d} (X - z) = \prod_{d|n} \phi_d(X) \quad (19.51)$$

où  $U_n$  est défini en 19.2.

*Démonstration.* En ce qui concerne la première égalité, tous les éléments de  $U_n$  sont des racines simples de  $X^n - 1$ . Donc le théorème 3.118 dit qu'il existe un nombre  $k$  (polynôme de degré zéro) tel que  $X^n - 1 = k \prod_{z \in U_n} (X - z)$ . Vu le coefficient du terme de plus haut degré, ce  $k$  ne peut être que 1.

Pour la suite nous utilisons l'union disjointe  $U_n = \bigcup_{d|n} \Delta_d$  de la proposition 19.13 et la définition (19.47) des polynômes cyclotomiques.  $\square$

#### Remarque 19.21.

Notons juste pour le plaisir que dans le produit  $\prod_{d|n} \prod_{z \in \Delta_d}$ , il y a bien  $n$  termes parce que  $\text{Card}(\Delta_d) = \varphi(d)$  et  $\sum_{d|n} \varphi(d) = n$  (définition 19.12 et proposition 5.38).

#### Proposition 19.22.

Les polynômes cyclotomiques sont à coefficients entiers :  $\phi_n \in \mathbb{Z}[X]$ .

*Démonstration.* Nous devons démontrer que les coefficients de  $\phi_n$  sont dans  $\mathbb{Z}$  alors qu'ils sont à priori dans  $\mathbb{C}$ . Nous démontrons cela par récurrence. D'abord  $\phi_1(X) = X - 1$ , d'accord. Ensuite

$$X^{n+1} - 1 = \prod_{d|n+1} \phi_d(X) = \phi_{n+1}(X) \cdot \underbrace{\prod_{\substack{d|n+1 \\ d \leq n}} \phi_d(X)}_{\in \mathbb{Z}[X] \text{ par récurrence}} \quad (19.52)$$

Le lemme 6.49 conclut que  $\phi_{n+1} \in \mathbb{Z}[X]$ . Nous avons considéré  $\mathbb{Z}$  comme sous anneau du corps  $\mathbb{C}$ .  $\square$

---

13. Définie par l'équation 19.24.

PropUImYnL

**Proposition 19.23.**

Soient  $1 \leq m \leq n$ , deux entiers, et

$$T(X) = \frac{X^n - 1}{X^m - 1} \in \mathbb{Z}(X). \quad (19.53)$$

Alors :

ItemhpDPKE

- (1) si  $m \mid n$  alors  $T \in \mathbb{Z}[X]$ ,
- (2) si  $m \mid n$  et si  $m < n$  alors  $\phi_n$  divise  $T$  dans  $\mathbb{Z}[X]$ .

*Démonstration.* Nous prouvons point par point.

- (1) Si  $m$  divise  $n$  alors les diviseurs de  $n$  sont l'union des diviseurs de  $m$  et des diviseurs de  $n$  qui ne divisent pas  $m$ . Soit

$$Q = \{\text{diviseurs de } n \text{ ne divisant pas } m\}. \quad (19.54)$$

Nous avons alors

$$X^n - 1 = \prod_{d \mid n} \phi_d(X) = \prod_{d \mid m} \phi_d(X) \cdot \prod_{q \in Q} \phi_q(X) = (X^m - 1) \cdot \prod_{q \in Q} \phi_q(X). \quad (19.55)$$

Nous avons donc

$$T(X) = \frac{X^n - 1}{X^m - 1} = \prod_{q \in Q} \phi_q(X) \in \mathbb{Z}[X]. \quad (19.56)$$

- (2) Nous venons de montrer que

$$T = \prod_{q \in Q} \phi_q \in \mathbb{Z}[X]. \quad (19.57)$$

Étant donné que  $m < n$  nous avons  $n \in Q$  et donc

$$T = \phi_n \cdot \prod_{q \in Q \setminus \{n\}} \phi_q. \quad (19.58)$$

Par conséquent  $\phi_n$  divise  $T$  dans  $\mathbb{Z}[X]$ .

□

CorTVUooErJiAC

**Corolaire 19.24.**

Si  $p$  est premier alors le polynôme cyclotomique  $\phi_p$  a une bonne tête :

$$\phi_p(X) = 1 + X + \cdots + X^{p-1}. \quad (19.59)$$

*Démonstration.* Nous utilisons la formule du lemme 19.20 en remarquant que seuls  $p$  et 1 divisent  $p$  :

$$X^p - 1 = \prod_{d \mid p} \phi_d(X) = \phi_1(X) \phi_p(X) = (X - 1) \phi_p(X). \quad (19.60)$$

Nous pouvons simplifier par  $X - 1$  en utilisant la formule du lemme 3.124(2) :

$$1 + X + \cdots + X^{p-1} = \phi_p(X) \quad (19.61)$$

□

PropoIeOVh

**Proposition 19.25** (Irréductibilité des polynômes cyclotomiques[630]).

Les polynômes cyclotomiques sont irréductibles sur  $\mathbb{Q}$ .

*Démonstration.* Pour rappel, nous savons déjà que pour tout  $n \in \mathbb{N}$ ,  $\phi_n \in \mathbb{Z}[X]$ . Puisque les racines de  $\phi_n$  sont les racines primitives de l'unité, nous devons montrer que toutes les racines primitives de l'unité ont même polynôme minimal (qui sera alors  $\phi_n$ ) ; en effet comme ces polynômes divisent  $\phi_n$ , s'ils sont distincts, la proposition 6.107 s'applique et le produit des polynômes minimaux diviserait  $\phi_n$ . Dans le cas inverse,  $\phi_n$  est polynôme minimal des racines primitives de l'unité et est donc irréductible. Soit donc  $\xi$ , une telle racine primitive. Une autre racine primitive est de la forme  $\xi^l$  où  $l$  est un nombre premier tel que  $\text{pgcd}(l, n) = 1$ .

Soient  $f$  et  $g$ , les polynômes minimaux dans  $\mathbb{Z}[X]$  de  $\xi$  et  $\xi^l$ . Nous allons montrer que  $f = g$  et donc que  $f = g = \phi_n$ . Supposons par l'absurde que  $f \neq g$ . Dans ce cas ils seraient des facteurs irréductibles distincts de  $\phi_n$  et il existerait un polynôme  $h$  tel que  $\phi_n = fgh$ . A priori,  $h \in \mathbb{Q}[X]$  parce que nous sommes justement en train de prouver que  $\phi_n$  est irréductible dans  $\mathbb{Q}[X]$ . Quoi qu'il en soit, le lemme de Gauss 6.51 nous montre que  $h \in \mathbb{Z}[X]$  parce que  $\phi_n$ ,  $f$  et  $g$  ont des coefficients entiers. Nous avons

$$f(\xi) = g(\xi^l) = 0. \quad (19.62)$$

Considérons le polynôme  $\psi(X) = g(X^l)$ . Ce polynôme  $\psi$  est dans  $\mathbb{Z}[X]$  et  $\psi$  est annulateur de  $\xi$ , donc  $f$  divise  $\psi$  en tant que polynôme minimal de  $\xi$ . Il y a un polynôme unitaire à coefficients entiers (lemme de Gauss forever)  $k$  tel que

$$\psi = fk \quad (19.63)$$

Nous considérons maintenant les projections sur  $\mathbb{F}_l[X]$  : étant donné que  $\phi_n = fgh$ , nous savons que  $\bar{f}\bar{g}$  divise  $\bar{\phi}_n$ . En même temps,  $\bar{f}$  divise  $\bar{\psi}$ . En utilisant le morphisme de Frobenius (c'est ici que la projection sur  $\mathbb{F}_l$  joue), nous avons aussi

$$\bar{\psi}(X) = \bar{g}(X^l) = \bar{g}(X)^l. \quad (19.64)$$

Par conséquent dire que  $\bar{f}$  divise  $\bar{\psi}$  revient à dire que  $\bar{f}(X)$  divise  $\bar{g}(X)^l$ . En particulier tout facteur irréductible de  $\bar{f}$  divise  $\bar{g}$ . Un facteur irréductible de  $\bar{f}$  serait donc à la fois dans  $\bar{f}$  et dans  $\bar{g}$  et donc deux fois (au moins) dans  $\bar{\phi}_n$  parce que  $\bar{f}\bar{g}$  divise  $\bar{\phi}_n$ . Dans un corps de décomposition de ce facteur,  $\phi_n$  aurait une racine double, alors que ce n'est pas le cas. Contradiction. Nous concluons que  $f = g$ .  $\square$

Le corolaire suivant va être utilisé pour déterminer les polygones constructibles à la règle et au compas, théorème de Gauss-Wantzel 19.89.

CorKRTooTJtyvP

### Corolaire 19.26.

Soit  $p$  un nombre premier et  $\alpha$  un entier non nul. Nous posons  $q = p^\alpha$ . Alors le polynôme minimal de  $e^{2i\pi/q}$  sur  $\mathbb{Q}$  est le polynôme cyclotomique  $\phi_q$ .

*Démonstration.* Le polynôme  $\phi_q$  est irréductible par la proposition 19.25, il est unitaire par définition et contient le monôme  $X - e^{2i\pi/q}$ , donc il est annulateur. Annulateur, irréductible et unitaire, la proposition 6.82(2) en fait le polynôme minimal de  $e^{2i\pi/q}$ .  $\square$

ThojCJpFW

### Théorème 19.27.

Soit  $P \in \mathbb{Z}[X]$  un polynôme unitaire irréductible non constant tel que toutes les racines dans  $\mathbb{C}$  soient de module  $\leq 1$ . Alors, soit  $P = X$ , soit  $P$  est un polynôme cyclotomique.

*Démonstration.* Le polynôme  $P = X$  vérifie les conditions. Pour la suite, nous supposons que  $P \neq X$ .

Nous notons  $P = \sum_i a_i X^i$ . Étant donné que  $P$  est irréductible et différent de  $X$ , nous avons  $a_0 \neq 0$  (sinon  $x = 0$  serait une racine). Nous allons montrer que les racines de  $P$  sont toutes des racines  $N$ -ièmes de l'unité (avec le même  $N$  pour toutes).

Soient  $\{\xi_i\}_{i=1,\dots,d}$  les racines de  $P$  ; on a

$$P = \prod_{i=1}^d (X - \xi_i) \quad (19.65)$$

avec  $\prod_{i=1}^d \xi_i = a_0$ . Par hypothèse,  $|\xi_i| \leq 1$  et donc  $0 < |a_0| \leq 1$ . Puisque  $P \in \mathbb{Z}[X]$  nous avons donc  $a_0 = 1$  et  $|\xi_i| = 1$ , pour tout  $i$ .

Nous introduisons les polynômes

$$g_q(X) = \prod_{i=1}^d (X - (\xi_i)^q), \quad (19.66)$$

et en particulier  $g_1 = P$ , et nous développons

$$g_q(X) = X^n + C_{1,q}X^{n-1} + \cdots + C_{n,q} \quad (19.67)$$

où

$$C_{k,q} = (-1)^k \sum_{1 \leq i_1 < \cdots < i_k \leq d} (\xi_{i_1} \cdots \xi_{i_k})^q. \quad (19.68)$$

Nous introduisons aussi les polynômes

$$F_{k,q}(X_1, \dots, X_n) = (-1)^k \sum_{1 \leq i_1 < \cdots < i_k \leq d} (X_{i_1} \cdots X_{i_k})^q \quad (19.69)$$

qui sont des polynômes symétriques. Ils vérifient deux propriétés. La première est que

$$C_{r,q} = F_{r,q}(\xi_1, \dots, \xi_n), \quad (19.70)$$

et la seconde est que les polynômes  $F_{r,1}$  sont les polynômes symétriques élémentaires à un coefficient près. Le théorème 6.180 nous donne alors des polynômes  $G_{k,q} \in \mathbb{Z}[X_1, \dots, X_n]$  tels que

$$F_{k,q}(X_1, \dots, X_n) = G_{k,q}(F_{1,1}(X_1, \dots, X_n), \dots, F_{k,1}(X_1, \dots, X_n)). \quad (19.71)$$

Nous savons que

$$|C_{k,q}| \leq \sum_{1 \leq i_1 < \cdots < i_k \leq d} 1 = \binom{d}{k}. \quad (19.72)$$

Donc  $g_q$  fait partie de l'ensemble fini des polynômes dans  $\mathbb{Z}[q]$  dont tous les coefficients sont bornés en valeur absolue par

$$\max_{k=1, \dots, d} \binom{d}{k}. \quad (19.73)$$

Il existe un certain nombre d'ensembles  $\{\xi_i\}$  qui sont racines de polynômes vérifiant les conditions du théorème. À chacun de ces ensembles est associé une suite de polynômes  $g_q$  et donc des coefficients  $C_{k,q}$ . Ce que nous avons vu est que l'ensemble de tous les coefficients  $C_{k,q}$  possibles (pour un choix donné des  $\{\xi_i\}$ ) est fini, en particulier, comme  $C_{1,q} = \sum_i \xi_i^q$ , pour chaque  $k$ , l'ensemble

$$\{\xi_k^q \text{ tel que } q \in \mathbb{N}\}. \quad (19.74)$$

Par le principe des tiroirs, il existe  $q_1$  et  $q_2$  tels que  $\xi_k^{q_1} = \xi_k^{q_2}$ . Ici,  $q_1$  et  $q_2$  dépendent de  $k$  et nous notons  $N_k = q_1 - q_2$ ; nous avons donc  $\xi_k^{N_k} = 1$ .

En posant  $N = \text{ppcm}(N_1, \dots, N_d)$ , nous avons

$$\xi_k^N = 1 \quad (19.75)$$

pour tout  $k$ .

Mais  $P$  est irréductible dans  $\mathbb{Z}[X]$ ; si il a  $\pm 1$  comme racines, alors c'est que  $P = X + 1$  ou  $P = X - 1$  et ce sont des polynômes cyclotomiques. Si  $P$  n'a pas  $\pm 1$  parmi ses racines, alors  $P$  n'a pas de racines dans  $\mathbb{Q}$  parce que  $\pm 1$  sont les seules racines de  $X^N - 1$  dans  $\mathbb{Q}$ .

Par conséquent  $P$  est un facteur irréductible de  $X^N - 1$  dans  $\mathbb{Q}[X]$ . Mais étant donné que

$$X^N - 1 = \prod_{d|N} \phi_d(X), \quad (19.76)$$

les polynômes cyclotomiques sont les seuls facteurs irréductibles de  $X^N - 1$ . Donc  $P$  est un polynôme cyclotomique.  $\square$

### 19.3.2 Nombres premiers

LemiAqLEn

**Lemme 19.28** ([631]).

Soit  $n \geq 1$ . Il existe un nombre premier  $p$  et un entier  $a$  tels que

- (1)  $p$  divise  $\phi_n(a)$ ,
- (2)  $p$  ne divise aucun de  $\phi_d(a)$  avec  $d \mid n$  et  $d \neq n$ .

De tels  $p$  et  $a$  vérifient automatiquement

- (1)  $p$  divise  $a^n - 1$ ,
- (2)  $p$  ne divise aucun des  $a^d - 1$  pour  $d \mid n$ ,  $d \neq n$ .

*Démonstration.* Nous posons

$$B(X) = \prod_{\substack{d \mid n \\ d \neq n}} \phi_d(X), \quad (19.77)$$

et nous commençons par montrer que  $\phi_n$  est premier avec  $B$ . Nous avons  $X^n - 1 = B\phi_n$ , donc  $B$  et  $\phi_n$  n'ont pas de racine commune (même pas dans  $\mathbb{C}$ ) parce que ce serait une racine double de  $X^n - 1$ . Notons que par définition 19.47, les polynômes cyclotomiques sont scindés (dans  $\mathbb{C}$ ), donc en particulier les polynômes  $\phi_n$  et  $B$  sont scindés, et donc premiers entre eux, dans  $\mathbb{C}$  et a fortiori dans  $\mathbb{Q}$ . Par Bézout (proposition 3.59), il existe  $U, V \in \mathbb{Q}[X]$  tels que

$$U\phi_n + VB = 1. \quad (19.78)$$

Si nous prenons  $a \in \mathbb{Z}$  tel que  $U' = aU$  et  $V' = aV$  soient tous deux dans  $\mathbb{Z}[X]$ , alors nous avons

$$U'\phi_n + V'B = a, \quad (19.79)$$

EqCaNMFi

égalité dans  $\mathbb{Z}[X]$ . Quitte à prendre un multiple assez grand de  $a$ , nous pouvons choisir  $a$  de telle sorte que  $|\phi_n(a)| \geq 2$ . Nous prenons alors un nombre premier  $p$  divisant  $\phi_n(a)$ .

Montrons que le  $a$  et le  $p$  ainsi construits satisfont aux exigences.

Puisque  $X^n - 1 = B\phi_n$ , si  $p$  divise  $\phi_n(a)$ , il divise automatiquement  $a^n - 1$  et donc  $[a^n]_p = 1$ , ce qui signifie entre autres que  $a$  et  $p$  sont premiers entre eux. Évaluons l'équation (19.79) en  $a$  :

$$U'(a)\phi_n(a) + V'(a)B(a) = a. \quad (19.80)$$

Le nombre  $p$  ne divisant pas  $a$ , mais divisant  $\phi_n(a)$ , il ne peut pas diviser  $B(a)$  <sup>14</sup>. Étant donné que  $p$  ne divise pas  $B(a)$ , il ne divise aucun des  $\phi_d(a)$  avec  $d \mid n$  et  $d \neq n$ .

Nous passons maintenant à la seconde partie de la preuve. Nous supposons avoir  $a$  et  $p$  tels que  $p$  soit un nombre premier divisant  $\phi_n(a)$  et tels que  $p$  ne divise aucun des  $\phi_d(a)$  avec  $d \mid n$ ,  $d \neq n$ . Le fait de diviser  $\phi_n(a)$  entraîne le fait de diviser  $a^n - 1$  parce que  $\phi_n$  est un des facteurs de  $X^n - 1$ . Soit maintenant  $d \neq n$  divisant  $n$ ; nous avons

$$X^d - 1 = \prod_{d' \mid d} \phi_{d'}, \quad (19.81)$$

EqwTWcCu

et cela est une partie du produit

$$\prod_{\substack{d \mid n \\ d \neq n}} \phi_d. \quad (19.82)$$

Puisque  $p$  ne divise aucun des  $\phi_d(a)$  de ce dernier produit, a fortiori, il ne divise pas le produit 19.81, et donc pas  $a^d - 1$ . □

LemrZnmpG

**Lemme 19.29.**

Si  $n \geq 1$ , alors il existe un nombre premier dans  $[1]_n$ , c'est-à-dire un nombre premier de la forme  $1 + kn$  avec  $k \in \mathbb{N} \setminus \{0\}$ .

14. C'est pour pouvoir dire ça que l'on a choisi  $V' \in \mathbb{Z}[X]$  de telle sorte que  $V'(a)$  soit dans  $\mathbb{Z}$

*Démonstration.* Soient  $n \geq 1$  et  $p, a$  les nombres donnés par le lemme 19.28. Puisque  $p$  divise  $\phi_n(a)$ ,  $p$  divise  $a^n - 1$  et donc  $[a]_p$  a un ordre qui divise  $n$  dans  $(\mathbb{Z}/p\mathbb{Z}) \setminus \{0\}$  parce que  $[a]_p^n = [1]_p$ .

Prenons  $d \neq n$  divisant  $n$ . Nous savons que

$$a^d - 1 = \prod_{d' \mid d} \phi_{d'}(a). \quad (19.83)$$

Par construction de  $a$  et  $p$ , nous avons

$$[\phi_{d'}(a)]_p \neq 0 \quad (19.84)$$

Comme  $\mathbb{Z}/p\mathbb{Z}$  est intègre, le produit est également non nul, c'est-à-dire

$$\left[ \prod_{d' \mid d} \phi_{d'}(a) \right]_p \neq 0, \quad (19.85)$$

et donc  $[a]_p^a \neq 1$ . Nous avons donc montré que si  $d \neq n$  divise  $n$ , alors nous avons en même temps

$$[a]_p^n = 1 \quad (19.86)$$

et

$$[a]_p^d \neq 1. \quad (19.87)$$

Cela prouve que  $[a]_p$  est d'ordre exactement  $n$ . Oui, mais l'ordre de  $[a]_p$  doit diviser l'ordre du groupe  $\mathbb{Z}/p\mathbb{Z}$  qui est  $p - 1$ , donc  $n$  divise  $p - 1$  et nous écrivons  $p = kn + 1$  avec  $k$  entier.  $\square$

ThoxwTjcl

**Théorème 19.30** (Forme faible du théorème de Dirichlet [136]).

Pour tout  $n \geq 1$ , il existe une infinité de nombres premiers dans  $[1]_n$ .

*Démonstration.* Le lemme 19.29 nous donne déjà l'existence de nombres premiers dans  $[1]_n$ . Il faut maintenant voir qu'il y en a une infinité. Nous supposons qu'il y en ait seulement un nombre fini :  $p_1, \dots, p_r$ , et nous notons

$$N = np_1 \dots p_r. \quad (19.88)$$

Nous utilisons maintenant le lemme 19.29 avec ce  $N$ , c'est-à-dire qu'on a un nombre premier de la forme

$$p = 1 + kN = 1 + knp_1 \dots p_r. \quad (19.89)$$

C'est un nombre premier plus grand que tous les  $p_i$ , et de la forme  $1 + \lambda n$ . Cela contredit l'exhaustivité de la liste  $p_1, \dots, p_r$ .  $\square$

## 19.4 Corps finis

SecCorpsFinizkAcbs

Si vous cherchez des choses à propos de RSA, c'est à la section 19.2.

### 19.4.1 Théorème de Wedderburn

ThoMncIWA

**Théorème 19.31** (Théorème de Wedderburn[188]).

Tout corps fini est commutatif.

*Démonstration.* Soit  $\mathbb{K}$  un corps fini et  $Z$ , le centre de  $\mathbb{K}$ . Ce dernier est un corps fini et un sous-corps de  $\mathbb{K}$ . Si  $q = \text{Card}(Z)$  alors par le lemme 6.58 nous avons

$$\text{Card}(\mathbb{K}) = q^n \quad (19.90)$$

pour un certain  $n$ .

Nous supposons maintenant que  $\mathbb{K}$  est non commutatif. Dans ce cas  $Z \neq \mathbb{K}$  et nous avons  $n \geq 2$ . Nous considérons aussi

$$Z_x = \{a \in \mathbb{K} \text{ tel que } ax = xa\}. \quad (19.91)$$

Le centre  $Z$  est un sous-corps de  $Z_x$ , donc il existe  $d(x)$  tel que

$$\text{Card}(Z_x) = q^{d(x)}. \quad (19.92)$$

De la même manière,  $Z_x$  est un sous-corps de  $\mathbb{K}$ , donc il existe  $m(x)$  tel que

$$\text{Card}(\mathbb{K}) = \text{Card}(Z_x)^{m(x)}. \quad (19.93)$$

En mettant bout à bout, nous avons

$$q^n = \text{Card}(Z_x)^{m(x)} = q^{d(x)m(x)}, \quad (19.94)$$

et par conséquent  $n = d(x)m(x)$ . Le point important à retenir est que  $d(x)$  divise  $n$  pour tout  $x \in \mathbb{K}$ .

Nous considérons maintenant l'action adjointe du groupe  $\mathbb{K} \setminus \{0\}$  sur lui-même :

$$\varphi(k)x = kxk^{-1}. \quad (19.95)$$

Nous notons  $\mathcal{O}_x$  l'orbite de  $x \in \mathbb{K} \setminus \{0\}$  pour cette action, et  $\text{Fix}(x)$  son stabilisateur. Nous avons

$$Z_y = \text{Fix}(y) \cup \{0\} \quad (19.96)$$

parce que  $Z_y$  et  $\text{Fix}(y)$  ont les mêmes définitions, sauf que  $\text{Fix}(y)$  est dans  $\mathbb{K} \setminus \{0\}$  alors que  $Z_y$  est dans  $\mathbb{K}$ . Nous avons donc

$$\text{Card}(\text{Fix}(y)) = \text{Card}(Z_y) - 1 = q^{d(y)} - 1. \quad (19.97)$$

Nous avons équivalence des affirmations suivantes :

- $\text{Card}(\mathcal{O}_x) = 1$
- $\mathcal{O}_x = \{x\}$
- $\text{Fix}(x) = \mathbb{K} \setminus \{0\}$
- $x \in Z \setminus \{0\}$ .

Soient  $z_0, \dots, z_{q-1}$  les éléments de  $Z$  avec  $z_0 = 0$ . Ce sont les éléments qui auront une orbite réduite à un point. Les orbites qui coupent  $Z \setminus \{0\}$  sont

$$\{z_1\}, \dots, \{z_{q-1}\} \quad (19.98)$$

et il y en a  $q-1$ . Soient  $\mathcal{O}_{y_1}, \dots, \mathcal{O}_{y_r}$ , les autres orbites. Nous utilisons l'équation des classes (2.80) :

$$\text{Card}(\mathbb{K}^*) = \text{Card}(Z^*) + \sum_{i=1}^r \frac{\text{Card}(\mathbb{K}^*)}{\text{Card}(\text{Fix}(y_i))}, \quad (19.99)$$

mais  $\text{Card}(Z^*) = q-1$ ,  $\text{Card}(\mathbb{K}^*) = q^n - 1$  et  $\text{Card}(\text{Fix}(y_i)) = q^{d(y_i)} - 1$ , donc

$$q^n - 1 = (q-1) + \sum_{i=1}^r \frac{q^n - 1}{q^{d(y_i)} - 1}. \quad (19.100) \quad \text{EqBPDzE}$$

Nous considérons la fraction rationnelle

$$F(X) = (X^n - 1) - \sum_{i=1}^r \frac{X^n - 1}{X^{d(y_i)} - 1}. \quad (19.101) \quad \text{EqATGciu}$$

Étant donné que  $d(y_i)$  divise  $n$ , nous avons, contrairement aux apparences,  $F \in \mathbb{Z}[X]$ , par la proposition 19.23(1).

Nous pouvons exploiter un peu mieux la proposition 19.23 en remarquant que  $d(y_i) < n$  parce que sinon  $\text{Card}(Z_{y_i}) = \text{Card}(\mathbb{K})$ , ce qui signifierait que  $y_i \in Z$ , ce qui nous avons exclu. Par conséquent le polynôme cyclotomique  $\phi_n$  divise

$$\frac{X^n - 1}{X^{d(y_i)} - 1} \quad (19.102)$$

dans  $\mathbb{Z}[X]$ . Le polynôme cyclotomique  $\phi_n$  divise également  $X^n - 1$  et par conséquent  $\phi_n$  divise  $F$ . Il existe donc  $Q \in \mathbb{Z}[X]$  tel que  $F = Q\phi_n$ . En particulier en évaluant en  $q$  :

$$F(q) = Q(q)\phi_n(q) = q - 1. \quad (19.103)$$

En effet nous avons  $F(q) = q - 1$  par construction : comparer (19.100) avec (19.101). Évidemment  $q \neq 1$  parce que si  $q = 1$  alors  $\text{Card}(\mathbb{K}) = 1$  et le théorème est trivial. Par ailleurs  $Q(q)$  est un entier (parce que  $Q \in \mathbb{Z}[X]$  et  $q \in \mathbb{N}$ ) et  $Q(q) \neq 0$ , parce qu'à droite de (19.103) nous avons  $q - 1 \neq 0$ . Nous avons donc  $|Q(q)| \geq 1$  et donc

$$|\phi_n(q)| \leq q - 1. \quad (19.104)$$

Par définition du polynôme cyclotomique nous avons

$$|\phi_n(q)| = \prod_{z \in \Delta_n} |q - z|. \quad (19.105)$$

Étant donné que ce produit doit être inférieur à  $q - 1$ , au moins un des termes doit l'être : il existe  $z_0 \in \Delta_n$  tel que  $|z_0 - q| \leq q - 1$ . Étant donné que  $n \geq 2$  nous avons  $z_0 \neq 1$ .

Mais d'autre part, comme indiqué sur la figure 19.1, la distance entre  $z_0$  et  $q$  doit être strictement plus grande que  $q - 1$  parce que  $q - 1$  est le minimum de la distance entre le cercle trigonométrique et  $q$ , et n'est atteint qu'en  $z = 1$ .

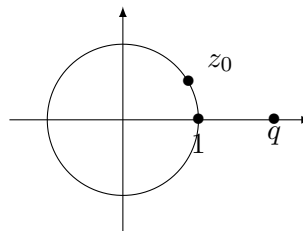


FIGURE 19.1: Nous devons avoir  $|z_0 - q| > q - 1$ .

LabelFigtrigoWedd

Nous avons ainsi obtenu une contradiction, et nous concluons que le corps  $\mathbb{K}$  est commutatif.  $\square$

### 19.4.2 Existence, unicité

Nous avons déjà défini le corps fini  $\mathbb{F}_p$  lorsque  $p$  est un nombre premier dans la section 6.1.3. Le théorème suivant sert à définir  $\mathbb{F}_{p^n}$  lorsque  $p$  est premier.

ThoOzgSfy

#### Théorème 19.32.

Soit  $p$  un nombre premier, soit  $n \in \mathbb{N} \setminus \{0\}$  et  $q = p^n$ . Alors il existe un unique corps  $\mathbb{K}$  de cardinal  $q$ . Ce corps est le corps de décomposition du polynôme  $X^q - X$  sur  $\mathbb{F}_p$ .

*Démonstration.* Montrons l'unicité. Soit  $\mathbb{K}$  un corps fini de cardinal  $q = p^n$ . Le groupe multiplicatif  $\mathbb{K}^*$  est de cardinal  $q - 1$ , et par le corolaire 2.15 tous les éléments de  $\mathbb{K}^*$  vérifient  $g^{q-1} = e$ , c'est-à-dire que dans  $\mathbb{K}[X]$ , les éléments de  $\mathbb{K}^*$  sont des racines du polynôme

$$X^{q-1} - 1 \quad (19.106)$$



Par conséquent  $\mathbb{K}$  est un corps de décomposition pour le polynôme  $Q(X) = X^q - X = X(X^{q-1} - 1)$  parce que  $Q(X) = 0$  dans  $\mathbb{K}$ . Il est unique par la proposition 6.145.

Montrons maintenant que le corps de décomposition de  $P = X^q - X$  sur  $\mathbb{F}_p$  est un corps de cardinal  $q$ . Pour ce faire nous considérons  $\mathbb{K}$  ce corps de décomposition, et  $\mathbb{E}$ , l'ensemble des racines de  $P$  dans  $\mathbb{K}$ . Nous allons montrer que  $\mathbb{E} = \mathbb{K}$  et que  $\mathbb{E}$  est un corps contenant  $q$  éléments.

Montrons que  $\mathbb{E}$  est un corps. Pour  $\alpha, \beta \in \mathbb{E}$  nous avons

$$(\alpha\beta)^q = \alpha^q \beta^q = \alpha\beta \quad (19.107)$$

parce que  $\alpha^q = \alpha$ . Le produit  $\alpha\beta$  est donc encore dans  $\mathbb{E}$ . Pour la somme,

$$(\alpha + \beta)^q = (\alpha + \beta)^{p^n} = \left((\alpha + \beta)^p\right)^{p^{n-1}} = (\alpha^p + \beta^p)^{p^{n-1}} = \dots = \alpha^{p^n} + \beta^{p^n} = \alpha + \beta. \quad (19.108)$$

En ce qui concerne l'inverse,

$$(\alpha^{-1})^q = (\alpha^q)^{-1} = \alpha^{-1}. \quad (19.109)$$

Donc  $\mathbb{E}$  est un corps. Évidemment  $\mathbb{E}$  est un corps de décomposition de  $P$  au sens où  $\mathbb{E}$  est une extension de  $\mathbb{F}_p$  sur lequel  $P$  est scindé (parce qu'il est scindé sur  $\mathbb{K}$  et  $\mathbb{E}$  est le sous-corps de  $\mathbb{K}$  contenant les racines de  $P$ ) et tel que  $\mathbb{E} = \mathbb{F}_p(\{\alpha_i\})$  où les  $\alpha_i$  sont les racines de  $P$ . Notons que  $\mathbb{F}_p \subset \mathbb{E}$  parce que dans  $\mathbb{F}_p$  on a  $x^q = x$ .

Par unicité, nous avons  $\mathbb{K} = \mathbb{E}$ . Nous devons montrer que  $P$  possède exactement  $q$  racines distinctes, afin d'avoir  $\text{Card}(\mathbb{E}) = q$ . Pour cela remarquons que

$$P'(X) = qX^{q-1} - 1 = -1 \quad (19.110)$$

dans  $\mathbb{F}_p$ . En effet  $P \in \mathbb{F}_p$  et  $q = 0$  dans  $\mathbb{F}_p$ . Par conséquent  $P'$  ne s'annule pas et  $P$  n'a pas de racine double. Toutes les racines étant simples, il y en a exactement  $q$ .  $\square$

Le théorème 19.32 ne permet pas de *construire* le corps à  $q = p^n$  éléments. Nous allons maintenant voir un certain nombre de résultats donnant des façons de le construire. Ces résultats proviennent de [632, 633, 634] et de [wikipedia](#)

Propnfebji

**Proposition 19.33** ([634]).

Soit  $\mathbb{K}$  un corps fini. Alors le groupe multiplicatif  $\mathbb{K} \setminus \{0\}$  est cyclique.

*Démonstration.* Soit  $\mathbb{K}$  un corps ayant  $q$  éléments. Le groupe  $\mathbb{K} \setminus \{0\}$  en a  $q - 1$ ; ergo l'ordre des éléments de  $\mathbb{K} \setminus \{0\}$  sont des diviseurs de  $q - 1$ ; c'est le corolaire 2.15. Soit  $d$  un diviseur de  $q - 1$  et

$$H_d \setminus \{0\} = \{x \text{ d'ordre } d \text{ dans } \mathbb{K} \setminus \{0\}\} \quad (19.111a)$$

$$H_d = \{\text{racines de } X^d - 1 \text{ dans } \mathbb{K}\}. \quad (19.111b)$$

Ici le polynôme  $X^d - 1$  est vu dans  $\mathbb{K}[X]$ . Notons que nous avons automatiquement  $H_d^* \subset H_d$ , mais l'inclusion inverse n'est pas assurée parce que les éléments d'ordre  $d/2$  par exemple sont aussi dans  $H_d$ . Supposons  $H_d^* \neq \emptyset$  et considérons  $a \in H_d^*$ . Alors l'application

$$\begin{aligned} \phi: \mathbb{Z}/d\mathbb{Z} &\rightarrow H_d \\ n &\mapsto a^n \end{aligned} \quad (19.112)$$

est un isomorphisme d'anneaux. En effet étant donné que  $a \in H_d^* \subset H_d$ , l'ensemble  $H_d$  contient le groupe cyclique engendré par  $a$ . Ce dernier contient, par construction,  $d$  éléments. Mais  $\text{Card}(H_d) \leq d$  parce que  $H_d$  est l'ensemble des racines d'un polynôme de degré  $d$ . Par conséquent  $\text{Card}(H_d) = d$  et l'ensemble  $H_d$  est bien engendré par  $a$  et  $\phi$  est bien un isomorphisme. Par conséquent tous les éléments de  $H_d^*$  sont des générateurs de  $H_d$ .

Inversement soit  $x$  un générateur de  $H_d$ . L'ordre de  $H_d$  étant  $d$ , l'ordre de  $x$  doit être un diviseur de  $d$ . Supposons donc que  $x$  soit d'ordre  $d/k$ . Dans ce cas nous devrions avoir  $\text{Card}(H_d) = d/k$ , ce qui contredit l'isomorphisme  $\phi$ .

En conclusion,  $H_d^*$  est l'ensemble des générateurs du groupe  $H_d$ . Le nombre de générateurs de  $\mathbb{Z}/d\mathbb{Z}$  étant  $\varphi(d)$  par la proposition 5.41, et  $H_d$  étant isomorphe à  $\mathbb{Z}/d\mathbb{Z}$  nous avons

$$\text{Card}(H_d^*) = \varphi(d). \quad (19.113)$$

Par conséquent si  $H_d^*$  n'est pas vide, son cardinal est  $\varphi(d)$ . Nous avons

$$q - 1 = \text{Card}(\mathbb{K}^*) \quad (19.114a)$$

$$= \text{Card} \left( \bigcup_{d|q-1} H_d^* \right) \quad (19.114b)$$

$$= \sum_{d|q-1} \text{Card}(H_d^*) \quad (19.114c)$$

$$\leq \sum_{d|q-1} \varphi(d) = q - 1 \quad (19.114d)$$

où nous avons utilisé la proposition 5.38. Par conséquent pour tout  $d$  divisant  $q - 1$  nous avons  $\text{Card}(H_d^*) = \varphi(d)$  et il y a au moins un élément d'ordre  $q - 1$  dans  $\mathbb{K}$ . Cet élément engendre  $\mathbb{K}^*$  parce que  $\mathbb{K}^*$  contient exactement  $q - 1$  éléments. Par conséquent  $\mathbb{K}^*$  est cyclique.  $\square$

CorpRUndR

### Corolaire 19.34.

Si  $p$  est un nombre premier, alors

$$(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}. \quad (19.115)$$

L'isomorphisme est un isomorphisme de groupes (abéliens). À gauche multiplicatif et à droite additif.

*Démonstration.* La proposition 19.33 nous enseigne que le groupe multiplicatif d'un corps fini est cyclique et donc isomorphe à un certain  $\mathbb{Z}/n\mathbb{Z}$ . Donc  $(\mathbb{Z}/p\mathbb{Z})^*$  est un groupe cyclique d'ordre  $p$ , et donc isomorphe à  $\mathbb{Z}/n\mathbb{Z}$  avec  $n = p - 1$ .  $\square$

Lorsque  $\mathbb{K}$  est un corps les éléments du groupe  $\mathbb{K}^*$  sont les **éléments primitifs** de  $\mathbb{K}$ .

propQRcU1q

### Proposition 19.35.

Soit  $\mathbb{K}$  un corps contenant  $q$  éléments. Alors

- (1)  $x^q = x$  pour tout  $x \in \mathbb{K}$ ,
- (2)  $X^q - X = \prod_{a \in \mathbb{K}} (X - a)$ .

*Démonstration.* Le groupe  $\mathbb{K} \setminus \{0\}$  ayant  $q - 1$  éléments, ses éléments vérifient  $a^{q-1} = 1$  par le corolaire 2.15 et par conséquent  $a^q = aa^{q-1} = a$ .

Soit  $a \in \mathbb{K}$ . Étant donné que  $a^q - a = 0$ , le polynôme  $(X - a)$  divise  $X^q - X$  dans  $\mathbb{K}[X]$ . Par conséquent

$$\prod_{a \in \mathbb{K}} (X - a) \quad (19.116)$$

divise également  $X^q - X$ . Les polynômes  $X^q - X$  et  $\prod_{a \in \mathbb{K}} (X - a)$  étant deux polynômes unitaires de même degré, le fait que l'un divise l'autre montre qu'ils sont égaux.  $\square$

Une conséquence de  $x^q = x$  est qu'il ne faut pas considérer le théorème 6.110 trop rapidement en disant « s'il s'annule partout, alors c'est le polynôme nul ». En effet dans un corps fini, « partout » n'est pas forcément très grand.

exVQB00BMPLkD

### Exemple 19.36.

Si  $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$  est le <sup>15</sup> corps à 3 éléments, alors le polynôme  $P(X) = X^3 - X$  s'évalue à zéro pour tout  $x \in \mathbb{F}_3$  (proposition 19.35.) mais il n'est pas le polynôme nul.  $\triangle$

15. Le singulier est justifié par le théorème 19.32, mais ça n'a pas d'importance ici.

### 19.4.3 Symboles de Legendre et carrés

Source : [635].

Nous disons que  $a \in \mathbb{F}_p$  est un **carré** si il existe  $b \in \mathbb{F}_p$  tel que  $a = b^2$ .

#### Définition 19.37.

Soit  $n \in \mathbb{N}$  et  $p > 2$  un nombre premier. Le **symbole de Legendre** est défini par

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & \text{si } p \text{ divise } n \\ 1 & \text{si } n \text{ est un carré dans } \mathbb{F}_p \\ -1 & \text{sinon.} \end{cases} \quad (19.117)$$

Note :  $-1$  peut être un carré, et pas que dans  $\mathbb{C}$ . Par exemple dans  $\mathbb{F}_5$  nous avons  $4 = -1$  et donc  $-1$  est un carré.

PropcGsJjk

#### Proposition 19.38.

Soit un nombre premier  $p > 2$ . Le corps  $\mathbb{F}_p^*$  contient autant de carrés que de non carrés. De plus pour tout  $n \in \mathbb{N}$  nous avons

$$\left(\frac{n}{p}\right) = n^{(p-1)/2} \pmod{p}. \quad (19.118)$$

Egbcugos

*Démonstration.* Nous considérons l'application

$$\begin{aligned} \psi: \mathbb{F}_p^* &\rightarrow \mathbb{F}_p^* \\ x &\mapsto x^2. \end{aligned} \quad (19.119)$$

C'est un morphisme de groupes multiplicatifs et  $\ker \psi = \{-1, 1\}$ . Étant donné que  $p > 2$ , nous avons alors

$$\text{Card}(\ker \psi) = 2 \quad (19.120)$$

parce que  $1 \neq -1$ . Évidemment l'ensemble des carrés dans  $\mathbb{F}_p^*$  est l'image de  $\psi$ . Le premier théorème d'isomorphisme 2.7(3) nous permet alors de conclure que

$$\text{Card}(\text{Image}(\psi)) = \frac{\text{Card}(\mathbb{F}_p^*)}{2}. \quad (19.121)$$

Ceci prouve la première assertion.

Par le petit théorème de Fermat (théorème 6.15), nous avons  $x^{p-1} = 1$  pour tout  $x \in \mathbb{F}_p^*$ . Les  $(p-1)$  éléments de  $\mathbb{F}_p^*$  sont donc tous racines d'un des deux polynômes

$$X^{(p-1)/2} = \pm 1. \quad (19.122)$$

Mais chacun des deux ne peut avoir, au maximum, que  $(p-1)/2$  solutions. Ils ont donc chacun exactement  $(p-1)/2$  racines.

Nous pouvons maintenant prouver la formule (19.118). D'abord si  $n = 0$ , elle est évidente. Si  $n$  est un carré dans  $\mathbb{F}_p$ , nous posons  $n = x^2$  et nous avons

$$n^{(p-1)/2} = x^{p-1} = 1 = \left(\frac{n}{p}\right). \quad (19.123)$$

Si  $n$  n'est pas un carré, c'est que  $n$  n'est pas une racine de  $X^{(p-1)/2} = 1$ . Le nombre  $n$  est alors une racine de  $X^{(p-1)/2} = -1$ . Nous avons alors

$$n^{(p-1)/2} = -1 = \left(\frac{n}{p}\right). \quad (19.124)$$

□

CoruJosNz

**Corolaire 19.39.**

Si  $a, b \in \mathbb{N}$  et si  $p > 2$  est un nombre premier, alors

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right) \left(\frac{b}{p}\right). \quad (19.125)$$

*Démonstration.* Par la formule (19.118),

$$\left(\frac{ab}{p}\right) = (ab)^{(p-1)/2} = a^{(p-1)/2} b^{(p-1)/2} = \left(\frac{a}{p}\right) \left(\frac{b}{p}\right). \quad (19.126)$$

□

Soit un nombre premier  $q > 2$  et  $\mathbb{A}$ , un anneau de caractéristique  $p$ . Si  $\alpha \in \mathbb{A}$  vérifie

$$1 + \alpha + \cdots + \alpha^{q-1} = 0, \quad (19.127)$$

nous définissons la **somme de Gauss** par

$$\tau = \sum_{x \in \mathbb{F}_q} \left(\frac{x}{q}\right) \alpha^x = \sum_{x=1}^{q-1} \left(\frac{x}{q}\right) \alpha^x. \quad (19.128)$$

Notons que la somme de Gauss dépend de  $q$  et du  $\alpha$  choisies.

PropciRUov

**Proposition 19.40.**

Les sommes de Gauss vérifient les propriétés suivantes.

- (1)  $\tau^2 = \left(\frac{-1}{q}\right) q$ . Nous allons noter  $\epsilon(q) = \left(\frac{-1}{q}\right)$ .
- (2) Si  $\mathbb{A}$  est de caractéristique  $p \geq 3$  et si  $p \neq q$  alors

$$\tau^p = \left(\frac{p}{q}\right) \tau. \quad (19.129) \quad \text{EqxBNpJz}$$

- (3) Si  $\mathbb{A}$  est de caractéristique  $p$  et si  $q$  est premier avec  $p$ , alors  $\tau$  est inversible dans  $\mathbb{A}$ .

*Démonstration.* D'abord nous notons que

$$\alpha^q - 1 = (\alpha - 1)(1 + \alpha + \cdots + \alpha^{q-1}) = 0 \quad (19.130)$$

par définition de  $\alpha$ . Nous calculons

$$\epsilon(q) \tau^2 = \epsilon(q) \sum_{x, y \in \mathbb{F}_q} \left(\frac{x}{q}\right) \left(\frac{y}{q}\right) \alpha^{x+y} \quad (19.131a)$$

$$= \sum_{x, y \in \mathbb{F}_q} \left(\frac{-xy}{q}\right) \alpha^{x+y}. \quad (19.131b) \quad \text{EqIObFco}$$

$$= \sum_{z \in \mathbb{F}_q} \sum_{y \in \mathbb{F}_q} \left(\frac{-(z-y)y}{q}\right) \alpha^z \quad (19.131c) \quad \text{EqWylhhk}$$

$$= \sum_{z \in \mathbb{F}_q} s_z \alpha^z \quad (19.131d) \quad \text{EqWolJszS}$$

Justifications :

- Pour obtenir (19.131b) nous avons utilisé le corolaire 19.39.
- (19.131c) est un changement de variable  $z = x + y$  dans la somme sur  $x$ .

— Pour (19.131d) nous avons posé

$$s_z = \sum_{y \in \mathbb{F}_q} \left( \frac{-(z-y)y}{q} \right). \quad (19.132)$$

Nous avons

$$s_0 = \sum_{y \in \mathbb{F}_q} \left( \frac{y^2}{q} \right). \quad (19.133)$$

Dans cette somme, tous les termes sont égaux à 1, sauf celui avec  $y = 0$  qui vaut zéro. Nous avons donc  $s_0 = q - 1$ . Voyons maintenant  $s_y$  avec  $y \neq 0$ . L'application

$$\begin{aligned} \mathbb{F}_q^* &\rightarrow \mathbb{F}_q \setminus \{1\} \\ k &\mapsto 1 - zy^{-1} \end{aligned} \quad (19.134)$$

étant une bijection nous pouvons effectuer le changement de variables  $t = y^{-1}z - 1$  pour la somme sur  $y$  en notant  $y^{-1}$  l'inverse de  $y$  dans  $\mathbb{F}_q^*$ , nous trouvons alors

$$\sum_{y \in \mathbb{F}_q} \left( \frac{y(z-y)}{q} \right) = \sum_{y \in \mathbb{F}_q} \left( \frac{y^2(y^{-1}z - 1)}{q} \right) \quad (19.135a)$$

$$= \sum_{y \in \mathbb{F}_q} \left( \frac{y^{-1}z - 1}{q} \right) \quad (19.135b)$$

$$= \sum_{t \in \mathbb{F}_q \setminus \{1\}} \left( \frac{t}{q} \right) \quad (19.135c)$$

$$= \underbrace{\sum_{t \in \mathbb{F}_q} \left( \frac{y}{q} \right)}_{=0} - \left( \frac{1}{1} \right) \quad (19.135d)$$

$$= -1 \quad (19.135e)$$

parce qu'il y a autant de carrés que de non carrés dans  $\mathbb{F}_q^*$  (proposition 19.38). En résumé nous avons

$$\epsilon(q)\tau^2 = \sum_{z \in \mathbb{F}_q} s_z \alpha^z \quad (19.136)$$

où

$$s_z = \begin{cases} q-1 & \text{si } z = 0 \\ -1 & \text{sinon.} \end{cases} \quad (19.137)$$

Cela donne

$$\epsilon(q)\tau^2 = (q-1) - \underbrace{(\alpha + \dots + \alpha^{q-1})}_{=-1} = q \quad (19.138)$$

où nous avons utilisé l'hypothèse sur  $\alpha$ . Donc  $\epsilon(q)\tau^2 = q$ , et étant donné que  $\epsilon(q) = \pm 1$  nous concluons

$$\tau^2 = \epsilon(q)q. \quad (19.139)$$

Nous prouvons maintenant la seconde partie. Comme  $\mathbb{A}$  est de caractéristique  $p$ , en utilisant le fait que le morphisme de Frobenius est un morphisme,

$$\tau^p = \left( \sum_{x \in \mathbb{F}_q} \left( \frac{x}{q} \right) \alpha^x \right)^p = \sum_{x \in \mathbb{F}_q} \left( \frac{x}{q} \right)^p \alpha^{px}. \quad (19.140)$$

Étant donné que  $\left( \frac{x}{q} \right) = \pm 1$  et que  $p$  est impair, nous avons

$$\left( \frac{x}{q} \right)^p = \left( \frac{x}{q} \right). \quad (19.141)$$

Du coup nous avons

$$\left(\frac{p}{q}\right) \tau^p = \sum_{x \in \mathbb{F}_p} \left(\frac{xp}{q}\right) \alpha^{px}. \quad (19.142)$$

Mais  $p$  étant inversible dans  $\mathbb{F}_q$ , l'application  $x \mapsto px$  est une bijection et nous pouvons sommer sur  $px$  au lieu de  $x$  :

$$\left(\frac{p}{q}\right) \tau^p = \sum_{x \in \mathbb{F}_p} \left(\frac{x}{q}\right) \alpha^x = \tau. \quad (19.143)$$

Nous trouvons alors que

$$\tau^p = \left(\frac{p}{q}\right) \tau. \quad (19.144)$$

Étant donné la formule du  $\tau^2$  que nous venons de démontrer, nous avons  $\tau^2 = \pm q$ . Les nombres  $p$  et  $q$  étant premiers entre eux, le théorème de Bézout (théorème 1.349) nous donne  $a$  et  $b$  tels que

$$ap + bq = 1. \quad (19.145)$$

Cela montre que  $b$  est un inverse de  $q$  modulo  $p$ . Donc  $\tau^2$  est inversible, et il en découle que  $\tau$  lui-même est inversible. □

ThoMiEiUm

**Théorème 19.41** (Loi de réciprocité quadratique).

Soient deux nombres premiers distincts  $p, q \geq 3$ . Alors

$$\left(\frac{p}{q}\right) = (-1)^{\frac{(p-1)(q-1)}{4}} \left(\frac{q}{p}\right). \quad (19.146)$$

*Démonstration.* Soit  $\phi_q$  le polynôme  $1 + X + \dots + X^{q-1}$  et l'anneau

$$\mathbb{A} = \mathbb{F}_p[X]/(\phi_q). \quad (19.147)$$

C'est un anneau de caractéristique  $p$  parce que son unité est le polynôme constant 1. Nous nommons  $\alpha = X/(\phi_q)$ , c'est-à-dire que  $\phi_q(\alpha) = 0$  dans  $\mathbb{A}$ , et nous pouvons considérer la somme de Gauss

$$\tau = \sum_{i \in \mathbb{F}_q} \left(\frac{i}{q}\right) \alpha^i. \quad (19.148)$$

Notons que ceci est un élément de  $\mathbb{A}$  et plus précisément un polynôme de degré zéro dans  $\mathbb{A}$ , et encore plus précisément, une classe d'un tel polynôme. Donc les coefficients de  $\alpha$  doivent être compris comme des éléments de  $\mathbb{F}_p$ . Nous savons (proposition 19.40) que

$$\tau^2 = \left(\frac{-1}{q}\right), \quad (19.149)$$

et en utilisant la formule (19.118) nous trouvons

$$\left(\frac{\tau^2}{p}\right) = (\tau^2)^{(p-1)/2} \mod p = \tau^{p-1} \mod p \quad (19.150)$$

En réalité sur cette dernière ligne, nous ne devrions pas préciser le « modulo  $p$  » parce que, comme mentionné plus haut, ce sont des éléments de  $\mathbb{F}_p$ . En utilisant cela, ainsi que (19.129), nous avons

$$\underbrace{\left(\frac{\tau^2}{p}\right)}_{\tau^{p-1}} \tau = \tau^p = \left(\frac{p}{q}\right) \tau \quad (19.151)$$

Puisque  $\tau$  est inversible, nous écrivons

$$\left(\frac{\tau^2}{p}\right) = \left(\frac{p}{q}\right) \quad (19.152)$$

Nous utilisons maintenant la formule (19.118) sur le membre de gauche avec  $n = \tau^2 = \left(\frac{-1}{q}\right)$  :

$$\left(\frac{p}{q}\right) = \left(\frac{-1}{q}\right)^{\frac{q-1}{2}} \left(\frac{q}{p}\right). \quad (19.153)$$

Toujours avec la même formule nous pouvons substituer  $\left(\frac{-1}{q}\right)$  par  $(-1)^{(q-1)/2}$  et obtenir

$$\left(\frac{p}{q}\right) = (-1)^{\frac{(q-1)}{2} \frac{(p-1)}{2}}. \quad (19.154)$$

□

Lemoabzrn

### Lemme 19.42.

Si  $p$  est un nombre premier  $p \geq 3$ , alors le symbole de Legendre  $x \mapsto \left(\frac{x}{p}\right)$  est l'unique morphisme non trivial de  $\mathbb{F}_p^*$  dans  $\{-1, 1\}$ .

*Démonstration.* Le fait que le symbole de Legendre soit non trivial est simplement le fait qu'il y ait des carrés et des non carrés dans  $\mathbb{F}_p^*$ ; voir la proposition 19.38. Pour l'unicité, soit  $\alpha: \mathbb{F}_p^* \rightarrow \{-1, 1\}$  un morphisme surjectif (c'est-à-dire non trivial). Étant donné que

$$\mathbb{F}_p^* = \ker(\alpha) \cup -\ker(\alpha), \quad (19.155)$$

le groupe  $\mathbb{F}_p^*/\ker(\alpha)$  ne contient que deux éléments :  $[1]$  et  $[-1]$ . Autrement dit,  $\ker(\alpha)$  est d'indice 2 dans  $\mathbb{F}_p^*$ .

Or  $\mathbb{F}_p \setminus \{0\}$  ne possède qu'un seul sous-groupe d'indice 2. En effet soit  $S$  un tel sous-groupe et  $a$ , un générateur de  $\mathbb{F}_p \setminus \{0\}$  (qui est cyclique par la proposition 19.33), alors  $a^2 \in S$  par le lemme 3.25. Par conséquent  $S$  contient le groupe des puissances paires de  $a$ . Le groupe  $S$  ne peut rien contenir de plus parce qu'il est d'indice 2 et que l'ordre de  $\mathbb{F}_p \setminus \{0\}$  est pair.

Bref, le sous-groupe  $\ker(\alpha)$  est l'unique sous-groupe d'indice 2 dans  $\mathbb{F}_p \setminus \{0\}$ . Mais la proposition 19.38 nous indique que  $|\ker(\alpha)| = \frac{p-1}{2}$ , c'est-à-dire que le groupe des carrés est d'indice 2. Nous avons donc, par l'unicité,

$$\ker(\alpha) = (\mathbb{F}_p \setminus \{0\})^2. \quad (19.156)$$

Au final, pour  $y \in \mathbb{F}_p \setminus \{0\}$ ,

$$\alpha(y) = \begin{cases} 1 & \text{si } y \text{ est un carré} \\ -1 & \text{sinon.} \end{cases} \quad (19.157)$$

Ce qui est bien la définition des symboles de Legendre. □

PROPooCLEUooHWnoXM

### Proposition 19.43.

Si  $p$  est un nombre premier, nous avons

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{si } p \in [1]_8 \text{ ou } p \in [7]_8 \\ -1 & \text{sinon.} \end{cases} \quad (19.158)$$

*Démonstration.* Soit le polynôme

$$X^4 + 1 \in \mathbb{F}_p[X] \quad (19.159)$$

et  $\alpha$ , une racine dans une extension<sup>16</sup> de  $\mathbb{F}_p$ <sup>17</sup>. Nous posons  $\theta = \alpha + \alpha^{-1}$  et nous calculons

$$\theta^2 = (\alpha + \alpha^{-1})(\alpha + \alpha^{-1}) = \alpha^2 + 2 + (\alpha^2)^{-1} = \alpha^2 + 2 - \alpha^2 = 2 \quad (19.160)$$

parce que  $\alpha^4$  étant  $-1$ , nous avons  $(\alpha^2)^{-1} = -\alpha^2$ . Bref,  $\theta^2 = 2$ .

16. Dans la source que je suivais (je ne sais plus où), on parlait ici de « fermeture » de  $\mathbb{F}_p$  et non d'extension. Il me semble que parler simplement d'extension suffit. Vous confirmez ?

17. Voir par exemple la proposition 6.142 pour l'existence d'une extension comme il faut.

Dire que 2 est un carré modulo  $p$  revient à dire que  $\theta$  est dans  $\mathbb{F}_p$ . C'est-à-dire que pour calculer le symbole de Legendre  $\left(\frac{2}{p}\right)$ , nous étudions pour quels  $p$ , l'élément  $\theta$  est vraiment dans  $\mathbb{F}_p$  et non seulement dans l'extension  $\mathbb{F}_p(\alpha)$ . En tenant compte de l'exemple 6.97, il faut distinguer deux cas :  $\alpha^p = \alpha$  et  $\alpha^p \neq \alpha$ . Autrement dit, si  $\alpha^k = \alpha$  pour un certain nombre premier  $k$ , alors le cas  $p = k$  est à traiter à part. La liste des puissances de  $\alpha$  est :

$$1, \alpha, \alpha^2, \alpha^3, -1, -\alpha, -\alpha^2, -\alpha^3, 1, \alpha, \dots \quad (19.161)$$

Nous avons donc automatiquement  $\alpha^{9k} = \alpha$ , mais  $p = 9k$  est exclu parce que  $p$  est premier. Nous devons donc vérifier si une des propriétés

$$\alpha^2 = \alpha \quad (19.162a)$$

$$\alpha^3 = \alpha \quad (19.162b)$$

$$-\alpha = \alpha \quad (19.162c)$$

$$-\alpha^3 = \alpha \quad (19.162d)$$

est possible. Il est aisément vérifiable, au cas par cas, que ces possibilités sont toutes incompatibles avec  $\alpha^4 = -1$ . Nous avons donc certainement  $\alpha^p \neq \alpha$  et compte tenu de l'exemple 6.97, l'équation  $x^p = x$  caractérise les éléments de  $\mathbb{F}_p$  dans  $\mathbb{F}_p(\alpha)$ .

L'équation  $X^2 = 2$  a exactement deux solutions qui sont  $\pm\theta$ . Nous avons donc  $2 \in \mathbb{F}_p^2$  si et seulement si  $\theta \in \mathbb{F}_p$  si et seulement si  $\theta^p = \theta$ . Nous avons réduit notre problème à déterminer pour quels  $p$  nous avons  $\theta^p = \theta$ . D'abord nous avons, par le morphisme de Frobenius,

$$\theta^p = (\alpha + \alpha^{-1})^p = \alpha^p + \alpha^{-p}. \quad (19.163)$$

Nous pouvons maintenant conclure facilement. Un nombre premier étant impair (sauf  $p = 2$  qui peut être traité à part),  $p$  est automatiquement dans un des ensembles  $[1]_8$ ,  $[3]_8$ ,  $[5]_8$  ou  $[7]_8$ . Nous avons quatre petites vérifications à faire. Dans tous les cas  $\alpha^{8k} = 1$ . Si  $p = 1 + 8k$ , alors

$$\theta^p = \alpha^{1+8k} + (\alpha^{-1})^{1+8k} = \alpha + \alpha^{-1} = \theta, \quad (19.164)$$

donc 2 est un carré dans  $\mathbb{F}_p$ . Si  $p \in [3]_8$ , alors  $\theta^p = \alpha^3 + \alpha^{-3}$ . Si cela était égal à  $\alpha + \alpha^{-1}$ , alors nous aurions

$$\alpha^6 + 1 = \alpha^4 + \alpha^2, \quad (19.165)$$

et donc  $\alpha^2 = 1$ , ce qui est impossible. Les vérifications pour  $p \in [5]_8$  et  $p \in [7]_8$  sont du même style.  $\square$

#### 19.4.4 Théorème de Chevalley-Warning

##### Lemme 19.44.

Soit  $\mathbb{K}$  un corps de caractéristique  $p$  et de cardinal  $q$ . Pour  $m \in \mathbb{N}$  nous définissons

$$S_m = \sum_{x \in \mathbb{K}} x^m. \quad (19.166)$$

Alors nous avons

$$S_m \mod p = \begin{cases} -1 & \text{si } m \geq 1 \text{ et } m \text{ divisible par } q-1 \\ 0 & \text{sinon.} \end{cases} \quad (19.167)$$

*Démonstration.* Si  $m = 0$ , alors  $x^0 = 1$  et  $S_m = q$ . Par conséquent  $S_m \mod p = 0$  parce que la caractéristique d'un corps divise son ordre (proposition 1.455).

Nous prenons maintenant  $m \geq 1$  et nous voyons séparément les cas où  $q-1$  divise  $m$  ou non. Si  $q-1$  divise  $m$ , alors pour tout  $x \neq 0$  nous avons

$$x^m = x^{k(q-1)} = 1 \quad (19.168)$$



parce que  $\mathbb{K} \setminus \{0\}$  est cyclique et  $x^{q-1} = 1$  par le petit théorème de Fermat (théorème 6.15). Par conséquent nous avons

$$\sum_{x \in \mathbb{K}} x^m = \sum_{x \in \mathbb{K} \setminus \{0\}} 1 = q - 1. \quad (19.169)$$

Si le nombre  $m \geq 1$  n'est pas divisible par  $q - 1$  alors nous prenons un générateur  $y$  du groupe  $\mathbb{K} \setminus \{0\}$ . Un tel élément vérifie  $y^m \neq 1$ . En effet, si  $y$  vérifiait  $y^m = 1$  alors cela signifierait que l'ordre de  $\mathbb{K} \setminus \{0\}$  est un diviseur de  $m$ , ce qui n'est pas le cas ici, parce que l'ordre de  $\mathbb{K} \setminus \{0\}$  est  $q - 1$ . Pour un tel  $y$ , l'application

$$\begin{aligned} \varphi: \mathbb{K} \setminus \{0\} &\rightarrow \mathbb{K} \setminus \{0\} \\ x &\mapsto yx \end{aligned} \quad (19.170)$$

est une bijection<sup>18</sup>. En ce qui concerne l'injectivité,  $ya = yb$  implique  $a = b$ . En ce qui concerne la surjectivité, si  $a$  est un générateur, si  $z = a^l$  et si  $y = a^k$ , alors

$$z = \varphi(a^{l-k}). \quad (19.171)$$

Nous pouvons maintenant poser le calcul.

$$S_m = \sum_{x \in \mathbb{K} \setminus \{0\}} x^m = \sum_{x \in \mathbb{K} \setminus \{0\}} (yx)^m = y^m \sum_{x \in \mathbb{K} \setminus \{0\}} x^m = y^m S_m. \quad (19.172)$$

Étant donné que  $y^m \neq 1$ , la seule solution est  $S_m = 0$ . □

ThoLTcYKk

**Théorème 19.45** (Chevalley-Warning[636]).

Soit  $\mathbb{K}$  un corps fini de cardinal  $q$  et de caractéristique  $p$ . Soient  $P_1, \dots, P_r$  des éléments de  $\mathbb{K}[X_1, \dots, X_n]$  tels que  $\sum_{i=1}^r \deg(P_i) < n$ . Nous considérons l'ensemble des zéros communs à tous les polynômes :

$$V = \{x \in \mathbb{K}^n \text{ tel que } P_1(x) = \dots = P_r(x) = 0\}. \quad (19.173)$$

Alors  $\text{Card}(V) \equiv 0 \pmod{p}$ .

*Démonstration.* Nous considérons le polynôme

$$P = \prod_{i=1}^r (1 - P_i^{q-1}). \quad (19.174)$$

Montrons que

$$P(x) = \begin{cases} 1 & \text{si } x \in V \\ 0 & \text{sinon.} \end{cases} \quad (19.175)$$

La première ligne est facile : étant donné que tous les  $P_i(x)$  sont nuls pour  $x \in V$ , nous avons  $P(x) = 1$ . Si  $x$  n'est pas dans  $V$ , alors nous avons un  $i$  tel que  $P_i(x) \in \mathbb{K} \setminus \{0\}$ . Mais dans ce cas (toujours la cyclicité de  $\mathbb{K} \setminus \{0\}$ ) nous avons  $P_i(x)^{q-1} = 1$  et donc le produit est nul.

En utilisant l'hypothèse sur le degré des  $P_i$ , nous trouvons

$$\deg(P) = \sum_{i=1}^r (q-1) \deg(P_i) < n(q-1). \quad (19.176)$$

Pour un polynôme  $Q \in \mathbb{K}[X_1, \dots, X_n]$ , nous définissons

$$\int Q = \sum_{x \in \mathbb{K}^n} Q(x). \quad (19.177)$$

---

18. Notons que nous n'avons pas réellement besoin que  $y$  soit un générateur. Nous n'utilisons seulement le fait que  $y^m \neq 1$  et  $y \neq 0$ .

Nous avons immédiatement

$$\int P = \sum_{x \in \mathbb{K}^n} P(x) = \sum_{x \in V} 1 = \text{Card}(V) \pmod{p}. \quad (19.178)$$

Nous insistons sur le « modulo  $p$  » parce que dans la formule  $P(x) = 1$ , le membre de droite est le 1 de  $\mathbb{K}$ ; il est donc automatiquement modulo la caractéristique de  $\mathbb{K}$ .

Il nous reste à prouver que  $\int P = 0$ . Pour cela nous décomposons

$$P = \sum_m c_m X_1^{m_1} \dots X_n^{m_n} \quad (19.179)$$

où la somme s'étend sur les  $m \in \mathbb{N}^n$  tels que  $c_m \neq 0$ . Nous avons

$$\int P = \sum_{x \in \mathbb{K}^n} \sum_m c_m x_1^{m_1} \dots x_n^{m_n} \quad (19.180a)$$

$$= \sum_m c_m \left( \sum_{x \in \mathbb{K}^n} x_1^{m_1} \dots x_n^{m_n} \right) \quad (19.180b)$$

$$= \sum_m c_m S_{m_1} \dots S_{m_n}. \quad (19.180c)$$

Le terme de plus haut degré dans la décomposition (19.179) est celui du  $m$  tel que  $\sum_i m_i$  est le plus grand. Comme ce degré est plus petit que  $n(q-1)$ , pour chacun des  $m$  rentrant dans la somme, nous avons

$$\sum_{i=1}^n m_i < n(q-1). \quad (19.181)$$

En particulier pour tout  $m \in \mathbb{N}^n$ , il existe  $i$  tel que  $m_i < q-1$ , et dans ce cas  $S_{m_i} = 0$ . Donc tous les termes de la somme

$$\sum_{m \in \mathbb{N}^n} c_m S_{m_1} \dots S_{m_n} \quad (19.182)$$

ont un facteur nul. □

CorfuHNKz

#### Corolaire 19.46.

Soit  $P_i$  des polynômes à  $n$  variables avec  $\sum_{i=1}^r \deg(P_i) < n$ . Si les  $P_i$  n'ont pas de terme constant, alors ils ont un zéro commun non trivial.

*Démonstration.* Nous reprenons les notations du théorème 19.45. Étant donné que les  $P_i$  n'ont pas de terme constant,  $0 \in V$ , mais  $\text{Card}(V) = 0 \pmod{p}$ . Par conséquent nous devons avoir  $\text{Card}(V) > p$ . □

#### Exemple 19.47.

Nous considérons les polynômes

$$P_1(x, y, t, u) = xy + x + ux \quad (19.183a)$$

$$P_2(x, y, t, u) = x + y - 3t. \quad (19.183b)$$

La somme de leurs degrés est 3 et ce sont des polynômes à 4 variables. Nous devons donc avoir, en vertu du corolaire 19.46, d'autres racines que la racine triviale  $(x, y, t, u) = (0, 0, 0, 0)$ .

Le corolaire nous donne aussi une borne inférieure du nombre de racines à chercher : plus que la caractéristique du corps sur lequel nous travaillons. Nous pouvons dire cela sans avoir la moindre idée de la façon dont on pourrait résoudre le système  $P_1 = P_2 = 0$ . △

### 19.4.5 Contenu d'un polynôme

LemHULrVaF

**Lemme 19.48** (de Gauss[126, 637]).

Soient  $P, Q \in \mathbb{Z}[X]$ . Alors

$$c(PQ) = c(P)c(Q) \quad (19.184)$$

où  $c$  est le contenu, définition 3.101.

*Démonstration.* Afin de fixer les notations, nous posons  $P = \sum_i a_i X^i$  et  $Q = \sum_j b_j Y^j$ .

(1) **Pour les polynômes primitifs**

Nous commençons par supposer que  $c(P) = c(Q) = 1$ . Dans ce cas si  $c(PQ) \neq 1$ , nous considérons un nombre premier  $p$  divisant  $c(PQ)$ . Puisque le contenu de  $P$  et de  $Q$  vaut 1, le nombre  $p$  ne peut pas diviser tous leurs coefficients. Nous définissons  $i_0$  de façon que  $a_{i_0}$  soit le premier à ne pas être divisible par  $p$ , et  $j_0$  de telle façon que  $b_{j_0}$  soit le premier à ne pas être divisible par  $p$ . Autrement dit :

$$p \mid a_0, p \mid a_1, \dots, p \mid a_{i_0-1}, p \nmid a_{i_0} \quad (19.185)$$

et de façon similaire pour  $j_0$ . Donc  $p$  ne divise ni  $a_{i_0}$ , ni  $b_{j_0}$ . Nous nous demandons alors avec malice quel est le coefficient de  $X^{i_0+j_0}$  dans  $PQ$ . La réponse est :

$$a_{i_0} b_{j_0} + \sum_{\substack{i+j=i_0+j_0 \\ i < i_0 \text{ ou } j < j_0}} a_i b_j. \quad (19.186)$$

Par définition  $p$  divise soit  $a_i$  soit  $b_j$  pour chacun des termes de la grande somme. Comme  $p$  ne divise pas  $a_{i_0} b_{j_0}$ , il ne divise pas le coefficient de  $X^{i_0+j_0}$  dans  $PQ$ , alors que nous étions partis en disant que  $p$  divisait tous les coefficients de  $PQ$ .

Nous concluons donc que  $c(PQ) = 1$ .

(2) **Cas général**

Si  $P$  et  $Q$  sont maintenant des polynômes sans condition particulière dans  $\mathbb{Z}[X]$ , nous considérons  $P_1 = \frac{P}{c(P)}$  et  $Q_1 = \frac{Q}{c(Q)}$  ; ces deux polynômes sont primitifs et nous avons alors, en utilisant la première partie :

$$c(P_1 Q_1) = 1. \quad (19.187)$$

Étant donné que

$$P_1 Q_1 = \frac{1}{c(P)c(Q)} PQ, \quad (19.188)$$

nous avons

$$c(PQ) = c(P)c(Q)c(P_1 Q_1) = c(P)c(Q). \quad (19.189)$$

□

### 19.4.6 Théorème de l'élément primitif

**Définition 19.49.**

Soit  $\mathbb{K}$  un corps. Une extension  $\mathbb{L}$  de  $\mathbb{K}$  est dite **finie** si  $\mathbb{L}$  est un espace vectoriel de dimension finie sur  $\mathbb{K}$ .

Notez que la définition d'extension finie ne suppose ni que  $\mathbb{K}$ , ni que  $\mathbb{L}$ , soient finis en tant qu'ensembles.

THOooYFLZooHYENKp

**Théorème 19.50** (de l'élément primitif). Si  $\mathbb{K}$  est un corps fini, toute extension finie de  $\mathbb{K}$  est simple<sup>19</sup>.

Si  $\mathbb{K}$  est un corps quelconque alors toute extension séparable finie est simple.

---

19. Définition 6.93.

*Démonstration.* Nous ne donnons la preuve que dans le cas où  $\mathbb{K}$  est fini. Dans ce cas nous savons par la proposition 19.33 que le groupe  $\mathbb{K} \setminus \{0\}$  est cyclique. Si de plus  $\mathbb{L}$  est une extension finie alors  $\mathbb{L}$  est fini en tant qu'ensemble. Par conséquent  $\mathbb{L} \setminus \{0\}$  est un groupe cyclique. Si  $\alpha$  est un générateur de  $\mathbb{L}$  alors  $\mathbb{L} = \mathbb{K}(\alpha)$  et l'extension est donc simple.

Une preuve de l'assertion dans le cas où  $\mathbb{K}$  est infini peut être trouvée sur wikipédia.  $\square$

PROPoWNHYooSzMJqF

**Proposition 19.51.**

*L'ordre d'un polynôme  $P$  vérifie les propriétés suivantes :*

- (1) *L'ordre de  $P$  est l'ordre multiplicatif de ses racines*
- (2) *L'ordre de  $P$  divise  $p^n - 1$ .*

LemZrUUOz

**Lemme 19.52.**

*Soit  $p$  un nombre premier et  $P$  un polynôme irréductible unitaire de degré  $n$ . Si  $\alpha, \beta \in \mathbb{F}_p[X]/P$ , alors  $(\alpha + \beta)^p = \alpha^p + \beta^p$ .*

*Démonstration.* La preuve est exactement la preuve classique<sup>20</sup> :

$$(\alpha + \beta)^p = \sum_k \binom{p}{k} \alpha^k \beta^{p-k} \quad (19.190)$$

où les coefficients binomiaux sont dans  $\mathbb{F}_p$  et donc nuls pour les  $k$  différents de  $p$  et de 0.  $\square$

Cette proposition est encore vraie avec  $\alpha, \beta \in \mathbb{F}_{p^n}$  et  $(\alpha + \beta)^{p^n}$ .

**Lemme 19.53.**

*Si  $\alpha \in \mathbb{F}_q$  est une racine d'ordre  $k$  de  $P$  (de degré  $n$ ) alors les racines de  $X^k - 1$  sont  $\{\alpha^i \text{ tel que } i = 0, \dots, k-1\}$ .*

Nous serions donc intéressés à construire  $\mathbb{F}_q$  comme quotient de  $\mathbb{F}_p[X]$  par un polynôme primitif. Le théorème suivant donne une description abstraite de  $\mathbb{F}_q$  qui va nous servir de point de départ pour la construction.

ThoqSludu

**Théorème 19.54** (Théorème de l'élément primitif).

*Soit  $p$  un nombre premier,  $n \in \mathbb{N}$  et  $q = p^n$ . Soit  $\mathbb{K}$  un corps à  $q$  éléments. Alors*

- (1) *Il existe  $\alpha \in \mathbb{K}$  tel que  $\mathbb{K} = \mathbb{F}_p[\alpha]$ .*
- (2) *Il existe un polynôme irréductible  $P \in \mathbb{F}_p[X]$  de degré  $n$  tel que*

$$\begin{aligned} \phi: \mathbb{F}_p[X]/(P) &\rightarrow \mathbb{K} \\ \overline{X} &\mapsto \alpha \end{aligned} \quad (19.191)$$

EgWIMhhm

*soit un isomorphisme de corps.*

*Soient  $\alpha$  et  $P$  choisis pour avoir les propriétés citées plus haut. Alors nous avons les propriétés suivantes.*

- (1)  *$P$  est primitif<sup>21</sup>.*
- (2)  *$P$  est scindé dans  $\mathbb{K}$ .*
- (3) *L'ensemble des racines de  $P$  est  $\{\alpha, \alpha^p, \dots, \alpha^{p^{n-1}}\}$ .*
- (4) *Le polynôme  $P$  divise  $X^q - X$  dans  $\mathbb{F}_p[X]$ .*

*Démonstration.* Le corps  $\mathbb{K}$  étant fini, il est cyclique par la proposition 19.33. Soit  $\alpha$  un générateur de  $\mathbb{K} \setminus \{0\}$  alors

$$\mathbb{K} = \mathbb{F}_p[\alpha]. \quad (19.192)$$

20. Voir les coefficients binomiaux de la proposition 1.375.

21. Définition 3.103.

Soit  $\ell$  le plus grand entier tel que l'ensemble

$$\{1, \alpha, \dots, \alpha^{\ell-1}\} \subset \mathbb{K} \quad (19.193)$$

soit libre. Pour rappel,  $\mathbb{K}$  est un espace vectoriel sur  $\mathbb{F}_p$ . Il existe des  $a_i \in \mathbb{F}_p$  tels que

$$\alpha^\ell + a_{\ell-1}\alpha^{\ell-1} + \dots + a_0 = 0. \quad (19.194)$$

De façon équivalente, il existe un polynôme unitaire  $P \in \mathbb{F}_p[X]$  de degré  $\ell$  tel que  $P(\alpha) = 0$ . Étant donné que  $\alpha$  est générateur de  $\mathbb{K}$ ,

$$\mathbb{K} = \text{Span}\{1, \alpha, \dots, \alpha^{\ell-1}\} \quad (19.195)$$

parce que  $\mathbb{K}$  est généré par les puissances de  $\alpha$  alors que les puissances de  $\alpha$  plus hautes que  $\ell - 1$  peuvent être générées par  $1, \alpha, \dots, \alpha^{\ell-1}$ . L'espace  $\mathbb{K}$  est donc un  $\mathbb{F}_p$ -espace vectoriel de dimension  $\ell$ ; par conséquent

$$\text{Card}(\mathbb{K}) = p^n = q \quad (19.196)$$

et  $\ell = n$ .

Montrons que  $P$  est irréductible dans  $\mathbb{F}_p$ . Si  $P$  était réductible dans  $\mathbb{F}_p$ , l'élément  $\alpha \in \mathbb{K}$  serait une racine d'un des facteurs, c'est-à-dire qu'il serait racine d'un polynôme de degré inférieur à  $n$ , ce qui contredirait le fait que

$$\{\alpha^{\ell-1}, \dots, 1\} \quad (19.197)$$

soit libre.

Montrons que l'application

$$\begin{aligned} \phi: \mathbb{F}_p[X]/(P) &\rightarrow \mathbb{K} \\ \overline{X} &\mapsto \alpha \end{aligned} \quad (19.198)$$

est un isomorphisme. Pour l'injectivité, deux éléments  $Q_1, Q_2 \in \mathbb{F}_p[X]/(P)$  s'écrivent

$$Q_1 = \sum_{k=0}^{n-1} a_k \overline{X}^k \quad (19.199a)$$

$$Q_2 = \sum_{k=0}^{n-1} b_k \overline{X}^k. \quad (19.199b)$$

Dans ce cas si  $\phi(Q_1) = \phi(Q_2)$  alors

$$\phi(Q_1) = \sum_{k=0}^{n-1} a_k \alpha^k = \phi(Q_2) = \sum_{k=0}^{n-1} b_k \alpha^k. \quad (19.200)$$

Mais l'ensemble  $\{1, \alpha, \dots, \alpha^{n-1}\}$  étant libre sur  $\mathbb{F}_p$ , cela implique  $a_k = b_k$ . La surjectivité de  $\phi$  provient du fait que  $\alpha$  génère  $\mathbb{K}$ .

Nous passons maintenant à la seconde partie de la démonstration. Soient  $\alpha \in \mathbb{K}$  tel que  $\mathbb{K} = \mathbb{F}_p[\alpha]$  et  $P \in \mathbb{F}_p[X]$  un polynôme irréductible de degré  $n$  tel que  $\alpha \mapsto \overline{X}$  soit un isomorphisme entre  $\mathbb{K}$  et  $\mathbb{F}_p[X]/(P)$ .

Le polynôme  $P$  est primitif parce que  $\alpha$  est d'ordre  $p^n$  dans  $\mathbb{K}$  alors que  $\overline{X} \mapsto \alpha$  est un isomorphisme. Par conséquent  $\overline{X}$  est d'ordre  $p^n$  dans  $\mathbb{F}_p[X]/P$ .

Nous commençons par prouver que l'ensemble

$$\{\alpha, \alpha^p, \alpha^{p^2}, \dots, \alpha^{p^{n-1}}\} \quad (19.201) \quad \text{EqAcsQHL}$$

est l'ensemble des racines distinctes de  $P$ . Pour cela nous posons

$$P(X) = \sum_{k=0}^n a_k X^k \quad (19.202)$$

avec  $a_k \in \mathbb{F}_p$ . D'abord  $\alpha$  est une racine de  $P$ . En effet

$$P(\overline{X}) = \sum_k a_k \overline{X}^k = 0 \quad \text{(19.203)}$$

parce que cette somme est calculée dans  $\mathbb{F}_p[X]/(P)$ . En appliquant l'isomorphisme  $\phi$  à l'égalité (19.203) nous trouvons

$$0 = \phi(P(\overline{X})) = \sum_k a_k \phi(\overline{X}^k) = \sum_k a_k \alpha^k. \quad (19.204)$$

Donc  $\alpha$  est bien une racine de  $P$  dans  $\mathbb{F}_p[X]$ . Nous devons montrer qu'il en est de même pour les autres puissances dans l'ensemble (19.201). Étant donné que pour tout  $x$  dans  $\mathbb{F}_p$  nous avons  $x^p = x$ , nous avons aussi

$$P(X^p) = \sum_k a_k (X^p)^k = \sum_k a_k^p (X^p)^k = \sum_k (a_k X^k)^p \quad (19.205)$$

alors que nous savons que  $x \mapsto x^p$  est un automorphisme de  $\mathbb{F}_p$  par la proposition 1.457. Par conséquent

$$P(X^p) = \sum_k (a_k X^k)^p = \left( \sum_k a_k X^k \right)^p = P(X)^p. \quad (19.206)$$

Nous avons montré que si  $\beta$  est une racine de  $P$ , alors  $\beta^p$  est également une racine de  $P$ . Nous savons déjà que  $\alpha$  est une racine de  $P$ , et que  $\alpha$  est également générateur de  $\mathbb{K}$ , c'est-à-dire que  $\alpha$  est d'ordre  $q - 1$ . Les puissances

$$\alpha, \alpha^p, \alpha^{p^2}, \dots, \alpha^{p^{n-1}} \quad (19.207)$$

sont donc distinctes ( $\alpha^{p^n} = \alpha^q = 1$ ) et sont toutes des racines de  $P$ . Étant donné que  $P$  est de degré  $n$  il ne peut pas y avoir d'autres racines. Nous concluons que l'ensemble

$$\{\alpha, \alpha^p, \alpha^{p^2}, \dots, \alpha^{p^{n-1}}\} \quad (19.208)$$

est l'ensemble des racines distinctes de  $P$  dans  $\mathbb{K}$ . Le polynôme  $P$  est alors scindé dans  $\mathbb{K}[X]$ .

Le dernier point du théorème est de montrer que  $P$  divise  $X^q - X$ . Pour cela nous allons montrer que toutes les racines de  $P$  sont des racines de  $X^q - X$ . Soit  $\beta$  une racine de  $P$ ; il s'écrit  $\beta = \alpha^k$  pour un certain  $k$ . Étant donné que  $\alpha^{q-1} = e = \alpha^{p^n-1}$ ,

$$\beta^q = (\alpha^{p^n})^k \quad (19.209a)$$

$$= (\alpha^{p^n-1} \alpha)^k \quad (19.209b)$$

$$= (\alpha^{q-1} \alpha)^k \quad (19.209c)$$

$$= \alpha^k \quad (19.209d)$$

$$= \beta. \quad (19.209e)$$

Cela signifie que  $\beta^q = \beta$  et donc que  $\beta$  est racine de  $X^q - X$ . □

### Corolaire 19.55.

*Le corps fini à  $q = p^n$  éléments est de caractéristique  $p$ .*

*Démonstration.* Nous considérons le corps fini  $\mathbb{K}$  à  $q$  éléments sous la forme  $\mathbb{K} = \mathbb{F}_p[X]/P$  comme indiqué par l'équation (19.191). Soit  $1_q$  la classe du polynôme 1 modulo  $P$ , nous considérons le morphisme

$$\begin{aligned} \mu: \mathbb{Z} &\rightarrow \mathbb{F}_q \\ n &\mapsto n1_q. \end{aligned} \quad (19.210)$$

Le noyau de cette application est  $\ker \mu = \mathbb{Z}_p$  parce que  $p1_q = 0$ , les coefficients étant à comprendre dans  $\mathbb{F}_p$ . □

DefnPNCFO

**Définition 19.56.**

Soient  $P$ , un polynôme de degré  $n$ , et  $p$ , un nombre premier. Un élément  $\alpha \in \mathbb{F}_p[X]/(P)$  est une **racine primitive** si les puissances de  $\alpha$  parcourent tout le groupe multiplicatif  $(\mathbb{F}_p[X]/P) \setminus \{0\}$ .

Lembccerei

**Lemme 19.57.**

Soit  $p$  un nombre premier et  $P$ , un polynôme de degré  $n$ . Si  $\alpha \in \mathbb{F}_p[X]/P$  est une racine primitive de  $P$  alors les autres racines de  $P$  sont également primitives.

*Démonstration.* Soit  $\alpha \in \mathbb{F}_p[X]/P$  une racine primitive de  $P$ . L'élément  $\alpha^p$  est également une racine parce que si  $P = \sum_k a_k X^k$ ,

$$P(\alpha^p) = \sum_k (a_k \alpha^k)^p = \left( \sum_k a_k \alpha^k \right)^p = 0 \quad (19.211)$$

où nous avons utilisé le fait que  $a_k^p = a_k$  étant donné que  $a_k \in \mathbb{F}_p$ . Par hypothèse  $\alpha$  est une racine primitive; cela implique que les éléments  $\alpha, \alpha^p, \alpha^{p^2}, \dots, \alpha^{p^{n-1}}$  sont distincts dans  $\mathbb{F}_p[X]/P$ . Ces éléments constituent donc toutes les racines de  $P$ .

Soit  $\beta = \alpha^{p^k}$  une racine de  $P$ . Montrons que  $\alpha$  est une puissance de  $\beta$ . Étant donné que  $(\mathbb{F}_p[X]/P) \setminus \{0\}$  est un groupe à  $p^n - 1$  éléments, le corolaire 2.15 indique que  $\alpha^{p^n} = \alpha$ . En particulier avec  $r = p^{n-k}$  nous avons

$$\beta^r = \alpha^{rp^k} = \alpha^{p^n} = \alpha. \quad (19.212)$$

Par suite toutes les puissances de  $\alpha$  sont des puissances de  $\beta$ , ce qui implique que  $\beta$  est générateur du groupe cyclique  $(\mathbb{F}_p[X]/P) \setminus \{0\}$ .  $\square$

LemkzWjse

**Lemme 19.58.**

Soit  $p$  un nombre premier et  $n$ , un entier. Un polynôme de degré  $d$ , irréductible dans  $\mathbb{F}_p[X]$ , divise  $X^{p^n} - X$  si et seulement si,  $d$  divise  $n$ .

**Théorème 19.59.**

Soient  $P$  et  $Q$  deux polynômes irréductibles de degré  $n$  dans  $\mathbb{F}_p[X]$ . Alors les quotients  $\mathbb{F}_p[X]/P$  et  $\mathbb{F}_p[X]/Q$  sont isomorphes en tant que corps.

En guise de démonstration de ce théorème, nous allons démontrer la proposition suivante.

PropCRPjZsp

**Proposition 19.60.**

Si  $\mathbb{K}$  et  $\mathbb{L}$  sont deux corps à  $q = p^n$  éléments, alors ils sont isomorphes.

*Démonstration.* Soit  $a$  un élément primitif de  $\mathbb{K}$  et  $P$  son polynôme minimal. Nous savons que  $\mathbb{K} \simeq \mathbb{F}_p[X]/P$  par le théorème de l'élément primitif 19.54. L'élément  $a$  est en particulier une racine de  $X^q - X$ . Par ailleurs  $P$  divise  $X^q - X$  par le lemme 19.58.

Nous avons aussi

$$X^q - X = \prod_{b \in \mathbb{L}} (X - b) \quad (19.213)$$

par la proposition 19.35. Étant donné que  $P$  divise  $X^q - X$ , un des éléments de  $\mathbb{L}$  annule  $P$ . Soit  $b \in \mathbb{L}$  tel que  $P(b) = 0$ . Soit  $Q$  le polynôme minimal de  $b$ . Par définition nous savons que  $Q$  divise  $P$ , mais  $P$  étant irréductible et unitaire, nous avons immédiatement  $P = Q$ . En particulier

$$\mathbb{F}_p[X]/P \simeq \mathbb{F}_p[X]/Q \simeq \mathbb{K}. \quad (19.214)$$

Nous montrons maintenant que  $\mathbb{F}_p[X]/Q \simeq \mathbb{L}$  par l'application

$$\begin{aligned} \phi: \mathbb{F}_p[X]/Q &\rightarrow \mathbb{L} \\ \overline{X} &\mapsto b \end{aligned} \quad (19.215)$$

qui se prolonge en  $R(\overline{X}) \mapsto R(b)$  pour tout  $R \in \mathbb{F}_p[X]$ . Cette application est bien définie parce que  $Q(b) = 0$ . Elle est injective parce que  $R(b) = 0$  ne peut pas avoir lieu avec  $R \in \mathbb{F}_p[X]/Q$  parce que  $Q$  est le polynôme minimal de  $b$ . La surjectivité vient alors du fait que les deux corps ont le même nombre d'éléments.  $\square$

### 19.4.7 Construction de $\mathbb{F}_{p^n}$

Le théorème 19.32 nous indique que, pour tout  $n \in \mathbb{N}$  et  $q = p^n$ , il existe un unique corps possédant  $q$  éléments. Ce corps est noté  $\mathbb{F}_q$ .

Le théorème 19.54 nous incite à chercher à écrire  $\mathbb{F}_q$  sous la forme

$$\mathbb{F}_q = \mathbb{F}_p[X]/(P) \quad (19.216)$$

pour un certain polynôme irréductible  $P \in \mathbb{F}_p[X]$ .

#### 19.4.7.1 La version du faignant

Nous pouvons construire le corps à  $q = p^n$  éléments en prenant le quotient de  $\mathbb{F}_p[X]$  par n'importe quel polynôme irréductible de degré  $n$ . Le résultat est le suivant.

PropHfrNCB

#### Proposition 19.61.

Soit  $P$  un polynôme unitaire irréductible dans  $\mathbb{F}_p[X]$ . Nous posons  $\mathbb{K} = \mathbb{F}_p[X]/(P)$ . Alors

- (1)  $\mathbb{K}$  est un corps à  $q$  éléments.
- (2)  $\alpha = \overline{X}$  est une racine de  $P$  dans  $\mathbb{K}$ .
- (3)  $\mathbb{K} = \mathbb{F}_p[\alpha]$ .

ItemiEFRTg

*Démonstration.* (1) En vertu du corolaire 6.44,  $\mathbb{K}$  est un corps. Il est aussi un espace vectoriel de dimension  $n$  sur  $\mathbb{F}_p$ , et contient donc  $p^n = q$  éléments.

(2) Nous avons  $P(\overline{X}) = 0$  par construction de  $\mathbb{K} = \mathbb{F}_p[X]/(P)$ .

(3) En tant que quotient de  $\mathbb{F}_p[X]$ , les éléments de  $\mathbb{K}$  sont des polynômes en  $\overline{X}$ .

□

#### 19.4.7.2 La version plus élaborée

Construire  $\mathbb{F}_q$  comme quotient de  $\mathbb{F}_p[X]$  par un polynôme irréductible quelconque ne donne pas d'information sur les générateurs de  $\mathbb{F}_q \setminus \{0\}$ , et en particulier il n'est pas toujours vrai que  $\overline{X}$  est générateur.

#### Exemple 19.62.

Construisons  $\mathbb{F}_4$ . Le polynôme  $X^2 + X + 1$  est irréductible dans  $\mathbb{F}_2$  parce qu'il n'a pas de racine (c'est vite vu : dans  $\mathbb{F}_2$  il n'y a que deux candidats). Donc  $\mathbb{F}_4 = \mathbb{F}_2[X]/(X^2 + X + 1)$ .  $\triangle$

#### Remarque 19.63.

Le corps  $\mathbb{F}_2$  n'est pas un sous-corps de  $\mathbb{C}$  parce que leurs caractéristiques ne sont pas les mêmes. Une conséquence est que les racines de polynômes peuvent être très différentes. Par exemple le polynôme  $X^2 + 1$  accepte  $x = 1$  comme racine dans  $\mathbb{F}_2$  tandis qu'il a pour racines  $\pm i$  dans  $\mathbb{C}$ .

En changeant de corps, les racines peuvent donc complètement changer. Ce n'est pas juste qu'il y a des racines dans l'un et pas dans l'autre.

ExemWUdrcs

#### Exemple 19.64.

Cherchons à construire  $\mathbb{F}_{16}$  comme quotient de  $\mathbb{F}_2$  par un polynôme de degré 4.

```
-----
| Sage Version 4.7.1, Release Date: 2011-08-11                |
| Type notebook() for the GUI, and license() for information.  |
|-----|
sage: x=polygen(GF(2))
sage: -x-1
x + 1
sage: Q=x**15-1
sage: Q.factor()
```



$$(x + 1) * (x^2 + x + 1) * (x^4 + x + 1) * (x^4 + x^3 + 1) \\ * (x^4 + x^3 + x^2 + x + 1)$$

Les polynômes candidats à avoir des racines génératrices sont donc au nombre de 3 :

$$P_1 = X^4 + X + 1 \quad (19.217a)$$

$$P_2 = X^4 + X^3 + 1 \quad (19.217b)$$

$$P_3 = X^4 + X^3 + X^2 + X + 1. \quad (19.217c)$$

Dans le quotient  $\mathbb{F}_2[X]/P_3$ , l'élément  $\bar{X}$  n'est pas générateur. En effet nous avons  $X^4 = X^3 + X^2 + X + 1$  et par conséquent les puissances successives de  $X$  sont

$$X \quad (19.218a)$$

$$X^2 \quad (19.218b)$$

$$X^3 \quad (19.218c)$$

$$X^4 = X^3 + X^2 + X + 1 \quad (19.218d)$$

$$1. \quad (19.218e)$$

La classe de  $X$  dans  $\mathbb{F}_2[X]/P_3$  n'est donc pas génératrice du groupe  $(\mathbb{F}_2[X]/P_3) \setminus \{0\}$ .

Le polynôme  $P_1 = X^4 + X + 1$  par contre est primitif parce que les puissances de  $X$  dans  $\mathbb{F}_2[X]/P_1$  sont

$$X \quad (19.219a)$$

$$X^2 \quad (19.219b)$$

$$X^3 \quad (19.219c)$$

$$X + 1 \quad (19.219d)$$

$$X^2 + X \quad (19.219e)$$

$$X^3 + X^2 \quad (19.219f)$$

$$X + 1 + X^3 \quad (19.219g)$$

$$X^2 + 1 \quad (19.219h)$$

$$X^3 + X \quad (19.219i)$$

$$X + 1 + X^2 \quad (19.219j)$$

$$X^2 + X + X^3 \quad (19.219k)$$

$$X^3 + X^2 + X + 1 \quad (19.219l)$$

$$1 + X^2 + X^2 \quad (19.219m)$$

$$1 + X^3 \quad (19.219n)$$

$$1 \quad (19.219o)$$

Cela fait 15 puissances distinctes, ce qui prouve que  $P_1$  est primitif. Nous verrons plus loin comment alléger un peu la vérification de la primitivité de  $P_1$ .  $\triangle$

PropNsLqWb

**Proposition 19.65** ([1]).

Soient un nombre premier  $p$ , un entier non nul  $n \in \mathbb{N} \setminus \{0\}$  ainsi qu'un polynôme  $P$  irréductible unitaire primitif dans  $\mathbb{F}_p[X]$ . Nous considérons  $\mathbb{K} = \mathbb{F}_p[X]/P$  et  $\alpha = \bar{X} \in \mathbb{K}$ . En notant  $q = p^n$  nous avons

(1) Les racines de  $P$  sont  $\{\alpha, \alpha^p, \dots, \alpha^{p^{n-1}}\}$  et  $\alpha^q = \alpha$ .

(2)  $P$  est le polynôme minimal de  $\alpha$ .

(3)  $P$  est scindé dans  $\mathbb{K}$ .

(4)  $P$  divise  $X^q - X$  dans  $\mathbb{K}$ .

(5) La famille  $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$  est une base de  $\mathbb{K}$  en tant qu'espace vectoriel sur  $\mathbb{F}_p$ .

(6) En tant qu'ensemble,

$$\mathbb{F}_q = \{0, \alpha, \alpha^2, \alpha^3, \dots, \alpha^{q-1}\}, \quad (19.220)$$

et les  $\alpha^k$  sont distincts pour  $k = 1, \dots, q-1$ .

*Démonstration.* La plupart des assertions sont des corolaires ou des paraphrases de résultats contenus dans les propositions précédentes.

- (1) L'assertion à propos des racines de  $P$  est contenue dans le lemme 19.57. D'autre part le groupe  $(\mathbb{F}_p[X]/P) \setminus \{0\}$  est cyclique d'ordre  $q-1$ . Par conséquent le corolaire 2.15 indique que  $\alpha^{q-1} = 1$  et donc  $\alpha^q = \alpha$ .
- (2) Soit  $\tilde{P}$  un polynôme annulateur de  $\alpha$ . Nous voyons que si  $\beta$  est racine de  $\tilde{P}$  alors  $\beta^p$  est également racine de  $\tilde{P}$  en utilisant les techniques habituelles. Par conséquent toutes les racines de  $P$  sont racines de  $\tilde{P}$ , ce qui implique que  $\tilde{P}$  est de degré au moins égal à celui de  $P$ .
- (3) Possédant  $n$  racines distinctes dans  $\mathbb{K}$ , le polynôme  $P$  est scindé.
- (4) D'après le lemme 19.35 un polynôme irréductible de degré  $n$  divise le polynôme  $X^{p^n} - X$ . Une autre façon de montrer ce point est de remarquer que le polynôme  $P$  est scindé et que toutes ses racines sont également racines de  $X^q - X$ .
- (5) Une combinaison linéaire nulle entre les éléments de  $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$  serait un polynôme annulateur de degré  $n-1$  de  $\alpha$ . Cet ensemble est donc libre. Par ailleurs un ensemble libre de  $n$  éléments dans un espace vectoriel de dimension  $n$  est générateur.
- (6) Si  $\alpha^l = \alpha^k$  avec  $k < l$  et  $k, l \leq q$  alors nous avons  $\alpha^r = 1$  avec  $r = l - k < q$ , ce qui contredirait la primitivité de  $P$ . Les éléments  $0, \alpha, \dots, \alpha^{q-1}$  étant distincts et au nombre de  $q$ , ils forment tout l'ensemble  $\mathbb{F}_q$ .

□

#### 19.4.8 Exemple : étude de $\mathbb{F}_{16}$

Dans cette section nous voulons construire  $\mathbb{F}_{16}$ . Nous considérons donc  $p = 2$  et  $n = 4$ . Des polynômes irréductibles de degré 4 dans  $\mathbb{F}_2[X]$  ne sont pas très difficiles à trouver. Par exemple  $X^4 + X^3 + X^2 + X + 1$ .

Si vous en voulez d'autres, en voici.

#### ⚠ Avertissement/question au lecteur !! 19.66

Le lemme suivant me semble douteux. Écrivez-moi si vous avez une preuve ou un contre-exemple.

LEMooTBROooANstIL

#### Lemme 19.67 ([1]).

Soit un polynôme de degré 4 dans  $\mathbb{F}_2[X]$ . Si il vérifie

- (1) le terme constant est non nul,
- (2) il y a un nombre impair de termes non nuls,

alors il est irréductible.

Les polynômes primitifs par contre, doivent être trouvés parmi les diviseurs irréductibles de  $X^{15} - 1$ . Montrons que

$$P = X^4 + X^3 + 1 \quad (19.221)$$

est primitif. Nous posons  $\omega = \overline{X} \in \mathbb{F}_2[X]/P$ . L'ordre de  $\omega$  dans le groupe  $(\mathbb{F}_2[X]/P) \setminus \{0\}$  doit être un diviseur de 15 et donc peut être seulement 1, 3, 5 ou 15. Le fait que l'ordre ne soit ni 1 ni 3 est trivial parce que le degré de  $P$  est 4. Montrons que l'ordre de  $\omega$  n'est pas 5 non plus :

$$\omega^5 = \omega^4 \omega = (\omega^3 + 1)\omega = \omega^4 + \omega = \omega^3 + \omega + 1 \neq 1. \quad (19.222)$$

Dans ce calcul nous avons abondamment utilisé le fait que  $-1 = 1$ .

À partir de maintenant nous posons  $\mathbb{K} = \mathbb{F}_2[X]/P$ . Les racines de  $P$  sont  $\omega, \omega^2, \omega^4$  et  $\omega^8$ . En effet si  $\beta$  est une racine de  $P$ , alors  $\beta^2$  est une racine en vertu de

$$P(\beta^2) = (\beta^2)^4 + (\beta^2)^3 + 1 = (\beta^4)^2 + (\beta^3)^2 + 1^2 = (\beta^4 + \beta^3 + 1)^2 = 0. \quad (19.223)$$

Ici nous avons implicitement utilisé le lemme 19.52. D'autre part  $P$  ne peut pas avoir plus de 4 racines.

**Proposition 19.68.**

*L'ensemble  $\{\omega, \omega^2, \omega^4, \omega^8\}$  est une base de  $\mathbb{F}_{16}$  sur  $\mathbb{F}_2$ .*

*Démonstration.* Nous savons que  $\{1, \omega, \omega^2, \omega^3\}$  est une base. En effet cet ensemble est libre (sinon  $\omega$  aurait un polynôme annulateur de degré 3) et générateur parce que l'espace engendré par 4 vecteurs indépendants sur  $\mathbb{F}_2$  contient  $2^4 = 16$  éléments.

Nous posons  $e_0 = 1$ ,  $e_1 = \omega$ ,  $e_2 = \omega^2$ ,  $e_3 = \omega^3$  et  $f_1 = \omega$ ,  $f_2 = \omega^2$ ,  $f_3 = \omega^4$ ,  $f_4 = \omega^8$ . En utilisant le calcul modulo  $\omega^4 + \omega^3 + 1 = 0$  et  $2 = 0$  nous trouvons

$$f_1 = \omega \quad (19.224a)$$

$$f_2 = \omega^2 \quad (19.224b)$$

$$f_3 = \omega^3 + 1 \quad (19.224c)$$

$$f_4 = \omega^3 + \omega^2 + \omega. \quad (19.224d)$$

Ensuite nous montrons que les vecteurs  $e_i$  peuvent être construits comme combinaisons linéaires des vecteurs  $f_j$  :

$$f_1 + f_2 + f_3 + f_4 = e_0 \quad (19.225a)$$

$$f_1 = e_1 \quad (19.225b)$$

$$f_2 = e_2 \quad (19.225c)$$

$$f_1 + f_2 + f_4 = e_3. \quad (19.225d)$$

Les quatre vecteurs  $f_j$  forment donc bien une base parce qu'ils sont générateurs d'un espace de dimension 4.  $\square$

**Exemple 19.69.** (1) Résoudre dans  $\mathbb{F}_{16}$  l'équation  $x^5 = a$  en discutant éventuellement en fonction de la valeur de  $a$ .

(2) Montrer qu'il existe quatre éléments  $\gamma \in \mathbb{F}_{16}$  tels que pour chacun d'eux l'ensemble  $B_\gamma = \{\gamma, \gamma^2, \gamma^4, \gamma^8\}$  est une base de  $\mathbb{F}_{16}$  sur  $\mathbb{F}_2$  telle que le produit de deux éléments de  $B_\gamma$  est, soit un élément de  $B_\gamma$ , soit 1.

C'est parti!

(1) Si  $a = 0$ , alors  $x = 0$  est la seule solution. Si  $a \neq 0$  alors  $a$  est une puissance de  $\omega$ ; nous posons  $a = \omega^l$ . Nous cherchons  $x$  sous la forme  $x = \omega^k$ . L'équation à résoudre pour  $k$  est

$$\omega^{5k} = \omega^l \quad (19.226)$$

où  $l$  est donné. Cette équation revient à

$$5k = l \pmod{15}. \quad (19.227)$$

Si  $l$  n'est pas un multiple de 5, alors il n'y a pas de solution. Il n'y a des solutions que pour  $l = 0, 5, 10$  et elles sont :

$$k = \begin{cases} 3, 6, 9, 12 & \text{si } l=0 \\ 1 & \text{si } l=5 \\ 2 & \text{si } l=10 \end{cases} \quad (19.228)$$

- (2) Nous cherchons  $\gamma$  sous la forme  $\gamma = \omega^k$ . Parmi les nombreuses contraintes liées à l'énoncé, nous devons avoir

$$\gamma^5 = 1, \gamma, \gamma^2, \gamma^4, \gamma^8. \quad (19.229)$$

Les possibilités  $\gamma^5 = \gamma, \gamma^2, \gamma^4, \gamma^8$  ne sont pas bonnes parce qu'elles impliqueraient que  $B_\gamma$  n'est pas une base. Reste à explorer  $\gamma^5 = 1$ .

Étant donné le premier point, nous restons avec les possibilités

$$\gamma = 1, \omega^3, \omega^6, \omega^9, \omega^{12}. \quad (19.230)$$

Évidemment  $\gamma = 1$  ne produit pas une base. Avec  $\gamma = \omega^3$  nous trouvons

$$B_\gamma = \{\omega^3, \omega^6, \omega^{12}, \omega^{24}\} = \{\omega^3, \omega^6, \omega^{12}, \omega^9\} \quad (19.231)$$

où nous avons utilisé le fait que  $\omega^k = \omega^{k \bmod 15}$ . En utilisant le fait que  $\omega^4 = \omega^3 + 1$  nous trouvons

$$\omega^5 = \omega^3 + \omega + 1 \quad (19.232a)$$

$$\omega^6 = \omega^3 + \omega^2 + \omega + 1 \quad (19.232b)$$

$$\omega^9 = \omega^2 + 1 \quad (19.232c)$$

$$\omega^{12} = \omega + 1. \quad (19.232d)$$

L'ensemble  $B_\gamma$  est alors formé des éléments

$$f_1 = \omega^3 \quad (19.233a)$$

$$f_2 = \omega^3 + \omega^2 + \omega + 1 \quad (19.233b)$$

$$f_3 = \omega + 1 \quad (19.233c)$$

$$f_4 = \omega^2 + 1. \quad (19.233d)$$

Il est assez simple de vérifier que c'est une base en remarquant que  $f_1 + f_2 + f_3 + f_4 = 1$ .

Les possibilités  $\gamma = \omega^6, \omega^9, \omega^{12}$  produisent les mêmes ensembles  $B_\gamma$ .

△

#### 19.4.9 Polynômes irréductibles sur $\mathbb{F}_q$

DefWXBkOxg

##### Définition 19.70.

La *fonction de Möbius* est la fonction  $\mu: \mathbb{N} \setminus \{0\} \rightarrow \{-1, 0, 1\}$  définie par

$$\mu(n) = \begin{cases} 0 & \text{si } n \text{ est divisible par un carré différent de } 1, \\ 1 & \text{si } n \text{ est le produit d'un nombre pair de nombres premiers distincts,} \\ -1 & \text{si } n \text{ est le produit d'un nombre impair de nombres premiers distincts,} \end{cases} \quad (19.234)$$

PROPooOVYJooFvmxyj

##### Proposition 19.71 ([638]).

Si  $m$  et  $n$  sont strictement positifs et premiers entre eux, alors

$$\mu(mn) = \mu(m)\mu(n). \quad (19.235)$$

De plus nous avons

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n > 1 \end{cases} \quad (19.236)$$

PropLBZoIoO

##### Proposition 19.72 (Formule d'inversion de Möbius[638]).

Soient  $f, g: \mathbb{N} \rightarrow \mathbb{C}$  telles que pour tout  $n \geq 1$ ,

$$g(n) = \sum_{d|n} f(d). \quad (19.237)$$

Alors

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) \quad (19.238)$$

où  $\mu$  est la fonction de Möbius pour tout  $n \geq 1$ .

LemRGuWqNu

**Lemme 19.73** ([639]).

Soient  $P, Q \in \mathbb{K}[X]$  ayant une racine commune dans une extension  $\mathbb{L}$  de  $\mathbb{K}$ . Si  $P$  est irréductible, alors  $P \mid Q$ .

*Démonstration.* Si  $P$  ne divise pas  $Q$ , alors  $P$  et  $Q$  sont premiers entre eux parce que dans la décomposition en irréductibles de  $Q$ , il n'y a pas de  $P$  tandis que dans celle de  $P$ , il n'y a que  $P$ . Par conséquent, il existe  $a, b \in \mathbb{K}[X] \subset \mathbb{L}[X]$  tels que <sup>22</sup>  $aP + bQ = 1$ . Cette dernière égalité est encore valable dans  $\mathbb{L}$  et donc rend impossible l'existence d'une racine commune.  $\square$

PropVFNOvzZ

**Proposition 19.74** ([639, 126]).

Soit  $p$  un nombre premier,  $n \geq 1$  et  $r \in \mathbb{N} \setminus \{0\}$ . Nous notons  $q = p^r$ ,  $A(n, q)$ , l'ensemble des polynômes unitaires irréductibles de degré  $n$  sur  $\mathbb{F}_q$ . Nous notons aussi  $I(n, q) = \text{Card}(A(n, q))$ . Alors :

(1) Le polynôme  $X^{q^n} - X$  se décompose en irréductibles de la façon suivante :

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in A(d, q)} P. \quad (19.239)$$

(2) Le nombre d'irréductibles est donné par

$$I(n, q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \quad (19.240)$$

où  $\mu$  est la fonction de Möbius (définition 19.70).

(3) Nous avons l'équivalence de suite

$$I(n, q) \sim_{n \rightarrow \infty} \frac{q^n}{n}. \quad (19.241)$$

*Démonstration.* (1) Soit un diviseur  $d$  de  $n$  et  $P \in A(d, q)$ . Montrons que  $P$  divise  $X^{q^n} - X$ . Nous considérons le corps  $\mathbb{K} = \mathbb{F}_q[X]/(P)$ , qui est une extension de degré  $\deg(P)$  de  $\mathbb{F}_q$  parce qu'il s'agit des polynômes de degré au maximum  $\deg(P)$  à coefficients dans  $\mathbb{F}_q$ . Ce corps possède donc  $q^d$  éléments et est isomorphe à  $\mathbb{F}_{q^d}$  par la proposition 19.60. Par construction dans  $\mathbb{K}$ , l'élément  $\alpha = [X]$  (la classe de  $X$  dans le quotient par  $P$ ) est une racine de  $P$ . Cet élément est également une racine de  $X^{q^d} - X$  parce que tout élément de  $\mathbb{F}_{q^d}$  est une racine de ce polynôme. Ce dernier point est la proposition 19.35.

Nous sommes donc dans la situation où  $P$  et  $X^{q^d} - X$  ont une racine commune dans l'extension  $\mathbb{F}_q[X]/(P)$ . Nous en déduisons que  $\alpha$  est aussi une racine de  $X^{q^n} - X$ . En effet en utilisant le fait que  $\alpha^{q^d} = \alpha$ , nous avons

$$\alpha^{q^n} = \alpha^{q^{kd}} = \alpha^{q^d q^{(k-1)d}} = \left(\alpha^{q^d}\right)^{q^{(k-1)d}} = \alpha^{q^{(k-1)d}}, \quad (19.242)$$

donc par récurrence, on a encore  $\alpha^{q^n} = \alpha$ , et  $\alpha$  est racine de  $X^{q^n} - X$ . Puisque  $P$  est irréductible, le lemme 19.73 nous indique alors, que  $P$  divise  $X^{q^n} - X$ .

Étant donné que tous les éléments de  $A(d, q)$  divisent  $X^{q^n} - X$  et sont irréductibles, leur produit divise encore  $X^{q^n} - X$  :

$$\prod_{d|n} \prod_{P \in A(d, q)} P \mid X^{q^n} - X. \quad (19.243)$$

22. Théorème de Bézout, 6.47.

Nous devons à présent montrer que tous les facteurs irréductibles de  $X^{q^n} - X$  sont dans un  $A(d, q)$  avec  $d \mid n$ . Soit donc  $P$  un facteur irréductible de  $X^{q^n} - X$  de degré  $d \geq 1$ . Nous posons encore  $\mathbb{K} = \mathbb{F}_q[X]/(P)$  et nous utilisons la propriété de multiplication sur les degrés (proposition 6.63) :

$$[\mathbb{F}_{q^n} : \mathbb{K}][\mathbb{K} : \mathbb{F}_q] = [\mathbb{F}_{q^n} : \mathbb{F}_q] = n, \quad (19.244)$$

donc  $[\mathbb{K} : \mathbb{F}_q]$ , qui vaut  $\deg(P)$  est un diviseur de  $n$ .

Étant donné que  $X^{q^n} - X$  n'a que des racines simples sur  $\mathbb{F}_{q^n}$  (à nouveau la proposition 19.35), dans sa décomposition en irréductibles sur  $\mathbb{F}_q$ , il n'a pas de facteur carré ; il n'a donc qu'une fois chacun des  $P \in A(d, q)$  avec  $d \mid n$ . Autrement dit, tous les facteurs irréductibles de  $X^{q^n} - X$  sont dans le produit  $\prod_{d \mid n} \prod_{P \in A(d, q)} P$  et donc  $X^{q^n} - X$  divise ce gros produit :

$$X^{q^n} - X \mid \prod_{d \mid n} \prod_{P \in A(d, q)} P. \quad (19.245)$$

Ayant déjà obtenu la divisibilité inverse et les polynômes étant unitaires, nous avons égalité.

(2) Nous passons au degré dans l'expression que nous venons de démontrer :

$$q^n = \sum_{d \mid n} d \operatorname{Card}(A(d, q)) = \sum_{d \mid n} dI(d, q). \quad (19.246)$$

Nous pouvons utiliser la formule d'inversion de Möbius (proposition 19.72) pour les fonctions  $g(n) = q^n$  et  $f(n) = dI(n, q)$ . Nous écrivons alors

$$f(n) = \sum_{d \mid n} \mu\left(\frac{n}{d}\right) q^d, \quad (19.247)$$

ou encore

$$I(n, q) = \frac{1}{n} \sum_{d \mid n} \mu\left(\frac{n}{d}\right) q^d, \quad (19.248)$$

ce qu'il fallait.

(3) Nous posons

$$r_n = \sum_{\substack{d \mid n \\ d < n}} \mu\left(\frac{n}{d}\right) q^d, \quad (19.249)$$

mais sachant que les diviseurs de  $n$ , outre  $n$  lui-même, sont tous plus petits ou égaux à  $n/2$  et qu'en valeur absolue, la fonction de Möbius est toujours plus petite ou égale à <sup>23</sup> 1,

$$|r_n| \leq \sum_{d=1}^{\lfloor n/2 \rfloor} q^d = \frac{q - q^{\lfloor n/2 \rfloor}}{1 - q} = q \frac{q^{\lfloor n/2 \rfloor} - 1}{q - 1} \leq \frac{q^{\lfloor n/2 \rfloor + 1}}{q - 1}. \quad (19.250)$$

D'autre part en reprenant la formule déjà prouvée,

$$I(n, q) = \frac{1}{n} \sum_{d \mid n} \mu\left(\frac{n}{d}\right) q^d = \frac{1}{n} \left( r_n + \mu\left(\frac{n}{n}\right) q^n \right) = \frac{r_n + q^n}{n}. \quad (19.251)$$

Au numérateur, le plus haut degré en  $n$  est  $q^n$  parce que  $r_n$  est en  $q^{\lfloor n/2 \rfloor}$ . Donc nous avons bien l'équivalence de suite pour  $n \rightarrow \infty$  :

$$\frac{q^n + r_n}{n} \sim_{n \rightarrow \infty} \frac{q^n}{n}. \quad (19.252)$$

□

23. Dans [126], ma dernière inégalité arrive comme une égalité.

### 19.4.10 Matrices

#### Proposition 19.75.

Nous avons

$$|\mathrm{GL}(n, \mathbb{F}_p)| = (p^n - 1)(p^n - p) \dots (p^n - p^{n-1}). \quad (19.253)$$

*Démonstration.* Par construction il existe une bijection entre  $\mathrm{GL}(n, \mathbb{F}_p)$  et l'ensemble des bases de  $\mathbb{F}_p^n$ . Nous devons donc seulement compter le nombre de bases. Pour le premier vecteur de base nous avons le choix entre les  $p^n - 1$  éléments non nuls de  $\mathbb{F}_p^n$ . Pour le second nous avons le choix entre  $p^n - p$  éléments, et ainsi de suite.  $\square$

LemcDOTzM

#### Lemme 19.76.

Soit  $\mathbb{K}$  un corps fini autre que  $\mathbb{F}_2$ <sup>24</sup>, soit un groupe abélien  $M$  et un morphisme  $\varphi: \mathrm{GL}(n, \mathbb{K}) \rightarrow M$ . Alors il existe un unique morphisme  $\delta: \mathbb{K} \setminus \{0\} \rightarrow M$  tel que  $\varphi = \delta \circ \det$ .

*Démonstration.* D'abord le groupe dérivé de  $\mathrm{GL}(n, \mathbb{K})$  est  $\mathrm{SL}(n, \mathbb{K})$  parce que les éléments de  $D(\mathrm{GL}(n, \mathbb{K}))$  sont de la forme  $ghg^{-1}h^{-1}$  dont le déterminant est 1.

De plus le groupe  $\mathrm{SL}(n, \mathbb{K})$  est normal dans  $\mathrm{GL}(n, \mathbb{K})$ . Par conséquent  $\mathrm{GL}(n, \mathbb{K})/\mathrm{SL}(n, \mathbb{K})$  est un groupe et nous pouvons définir l'application relevée

$$\tilde{\varphi}: \frac{\mathrm{GL}(n, \mathbb{K})}{\mathrm{SL}(n, \mathbb{K})} \rightarrow M \quad (19.254)$$

vérifiant  $\varphi = \tilde{\varphi} \circ \pi$  où  $\pi$  est la projection.

Nous pouvons faire la même chose avec l'application

$$\det: \mathrm{GL}(n, \mathbb{K}) \rightarrow \mathbb{K} \setminus \{0\} \quad (19.255)$$

qui est un morphisme de groupes dont le noyau est  $\mathrm{SL}(n, \mathbb{K})$ . Cela nous donne une application

$$\tilde{\det}: \frac{\mathrm{GL}(n, \mathbb{K})}{\mathrm{SL}(n, \mathbb{K})} \rightarrow \mathbb{K} \setminus \{0\} \quad (19.256)$$

telle que  $\det = \tilde{\det} \circ \pi$ . Cette application  $\tilde{\det}$  est un isomorphisme. En effet, elle est surjective parce que le déterminant l'est, et elle est injective parce que son noyau est précisément ce par quoi on prend le quotient. Par conséquent  $\tilde{\det}$  possède un inverse et nous pouvons écrire

$$\varphi = \tilde{\varphi} \circ \tilde{\det}^{-1} \circ \det \circ \pi. \quad (19.257)$$

Étant donné que  $\tilde{\det} \circ \pi = \det$ , nous avons alors  $\varphi = \delta \circ \det$  avec  $\delta = \tilde{\varphi} \circ \tilde{\det}^{-1}$ . Ceci conclut la partie existence de la preuve.

En ce qui concerne l'unicité, nous considérons  $\delta': \mathbb{K} \setminus \{0\} \rightarrow M$  telle que  $\varphi = \delta' \circ \det$ . Pour tout  $u \in \mathrm{GL}(n, \mathbb{K})$  nous avons  $\delta'(\det(u)) = \varphi(u) = \delta(\det(u))$ . L'application  $\det$  étant surjective depuis  $\mathrm{GL}(n, \mathbb{K})$  vers  $\mathbb{K} \setminus \{0\}$ , nous avons  $\delta' = \delta$ .  $\square$

#### Théorème 19.77.

Soient  $p \geq 3$  un nombre premier et  $E$ , un  $\mathbb{F}_p$ -espace vectoriel de dimension finie  $n$ . Pour tout  $u \in \mathrm{GL}(E)$  nous avons

$$\epsilon(u) = \left( \frac{\det(u)}{p} \right). \quad (19.258)$$

Ici  $\epsilon$  est la signature de  $u$  vue comme une permutation des éléments de  $\mathbb{F}_p$ .

*Démonstration.* Commençons par prouver que

$$\epsilon: \mathrm{GL}(E) \rightarrow \{-1, 1\}. \quad (19.259)$$

---

24. Je ne comprends pas très bien à quel moment joue cette hypothèse.

est un morphisme. Si nous notons  $\bar{u} \in S(E)$  l'élément du groupe symétrique correspondant à la matrice  $u \in \mathrm{GL}(E)$ , alors nous avons  $\overline{uv} = \bar{u} \circ \bar{v}$ , et la signature étant un morphisme (proposition 1.291),

$$\epsilon(uv) = \epsilon(\bar{u} \circ \bar{v}) = \epsilon(\bar{u})\epsilon(\bar{v}). \quad (19.260)$$

Par ailleurs  $\{-1, 1\}$  est abélien, donc le lemme 19.76 s'applique et nous pouvons considérer un morphisme  $\delta: \mathbb{F}_p \setminus \{0\} \rightarrow \{-1, 1\}$  tel que  $\epsilon = \delta \circ \det$ .

Nous allons utiliser le lemme 19.42 pour montrer que  $\delta$  est le symbole de Legendre. Pour cela il nous faudrait trouver un  $x \in \mathbb{F}_p \setminus \{0\}$  tel que  $\delta(x) = -1$ . Étant donné que  $\det$  est surjective, nous cherchons ce  $x$  sous la forme  $x = \det(u)$ . Par conséquent nous aurions

$$\delta(x) = (\delta \circ \det)(u) = \epsilon(u), \quad (19.261)$$

et notre problème revient à trouver une matrice  $u \in \mathrm{GL}(E)$  dont la permutation associée soit de signature  $-1$ .

Soit  $n = \dim E$ ; en conséquence de la proposition 19.61(3), l'espace  $\mathbb{E}_q = \mathbb{F}_{p^n}$  est un  $\mathbb{F}_p$ -espace vectoriel de dimension  $n$  et est donc isomorphe en tant qu'espace vectoriel à  $E$ . Étant donné que  $\mathbb{F}_q$  est un corps fini, nous savons que  $\mathbb{F}_q \setminus \{0\}$  est un groupe cyclique à  $q - 1$  éléments. Soit  $y$ , un générateur de  $\mathbb{F}_q \setminus \{0\}$  et l'application

$$\begin{aligned} \beta: \mathbb{F}_q &\rightarrow \mathbb{F}_q \\ x &\mapsto yx. \end{aligned} \quad (19.262)$$

Cela est manifestement  $\mathbb{F}_p$ -linéaire (ici  $y$  et  $x$  sont des classes de polynômes et  $\mathbb{F}_p$  est le corps des coefficients). L'application  $\beta$  fixe zéro et à part zéro, agit comme le cycle

$$(1, y, y^2, \dots, y^{q-2}). \quad (19.263)$$

Nous savons qu'un cycle de longueur  $n$  est de signature  $(-1)^{n+1}$ . Ici le cycle est de longueur  $q - 1$  qui est pair (parce que  $p \geq 3$ ) et par conséquent, l'application  $\beta$  est de signature  $-1$ .  $\square$

## 19.5 Constructions à la règle et au compas

**Définition 19.78** ([640]).

Soit  $E$  une partie de  $\mathbb{R}^2$ . Un point de  $\mathbb{R}^2$  est constructible en une étape à partir de  $E$  si il est un point de  $E$  ou une intersection de deux objets parmi

- les droites passant par deux points distincts de  $E$ ;
- les cercles centrés en un point de  $E$  et dont le rayon est la distance entre deux points de  $E$ .

Nous notons  $C_1(E)$  l'ensemble des points constructibles en une étape à partir de  $E$ .

Les points constructibles en  $n$  étapes à partir de  $E$  sont définis par récurrence :  $C_{n+1}(E) = C_1(C_n(E))$ . Enfin un point de  $\mathbb{R}^2$  est **constructible** à partir de  $E$  si il appartient à

$$C(E) = \bigcup_{n=1}^{\infty} C_n(E). \quad (19.264)$$

Un réel est **constructible** si il est l'abscisse d'un point constructible.

Pour toute la suite, nous allons considérer les points et réels constructibles à partir de l'ensemble  $E = \{(0, 0), (0, 1)\}$ .

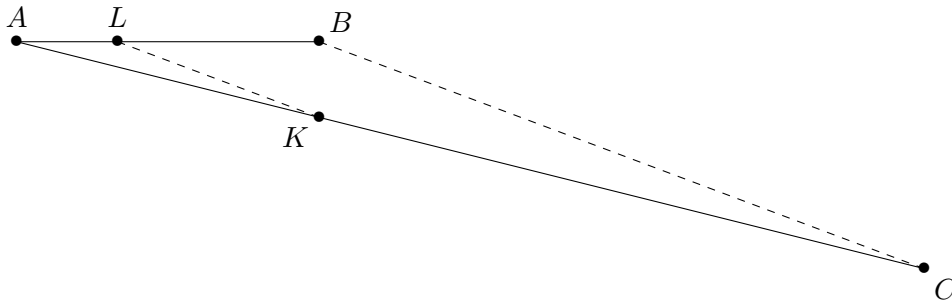
### 19.5.1 Quelques constructions

**Proposition 19.79** ([641]).

Les nombres rationnels sont tous constructibles.

*Démonstration.* Si le réel  $r$  est constructible, alors  $kr$  est également constructible pour tout  $k \in \mathbb{Z}$ . Nous devons donc seulement pouvoir construire le nombre  $1/n$  pour tout  $n \in \mathbb{N} \setminus \{0\}$ .





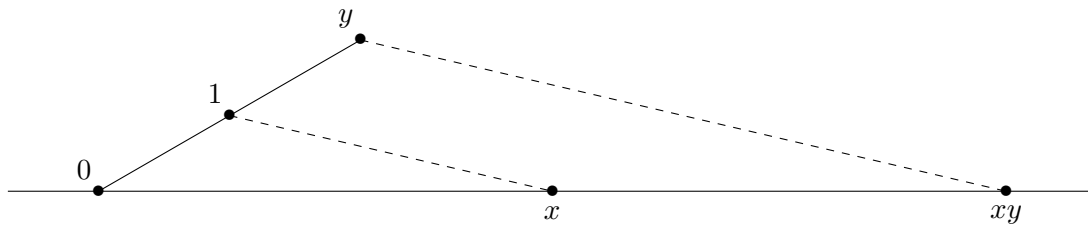
La méthode pour construire le nombre  $1/n$  est la suivante. Soit  $[AB]$  un segment de longueur 1 (par exemple  $A = (0, 0)$  et  $B = (1, 0)$ ) et un point  $C$ , non aligné avec  $A$  et  $B$ , et tel que  $[AC]$  ait une longueur  $n$ . Nous plaçons sur  $[AC]$  le point  $K$  situé à une distance 1 de  $A$  en pointant le compas en  $A$  et en traçant le cercle de rayon  $[AB]$ .

La droite passant par  $K$  et parallèle à  $(BC)$  coupe  $[AB]$  en un point  $L$ . Maintenant, le segment  $[AL]$  a une longueur de  $1/n$ , par le théorème de Thalès<sup>25</sup>.  $\square$

ExGROooIosiBt

**Exemple 19.80** (Multiplication à la règle et au compas<sup>[640]</sup>).

Soient  $x$  et  $y$  deux nombres constructibles. Montrons qu'il est possible de construire le nombre  $xy$ . La construction est la suivante :



- On trace deux droites sécantes en  $A$ .
- Sur la première nous plaçons le point  $Y$  à distance  $y$  de  $A$  et le point  $P$  à distance 1 de  $A$ .
- Sur la seconde on place le point  $X$  à distance  $x$  de  $A$ .
- On trace la droite  $(PX)$
- Puis la parallèle à  $(PX)$  passant par  $Y$ .
- Le point d'intersection entre cette dernière droite et  $(AX)$  est le point  $B$ .

La longueur  $AB$  est égale à  $xy$ .  $\triangle$

ExTYMooSMCvSr

**Exemple 19.81** (Racine carrée à la règle et au compas<sup>[640]</sup>).

Nous supposons que le nombre  $x$  est constructible, et nous voulons une construction qui donne un segment de longueur  $\sqrt{x}$ . Nous traçons un segment  $[BC]$  dont la longueur correspond à la plus grande des valeurs entre  $x$  et 1, puis le cercle de diamètre  $BC$ , ensuite le point  $H$  sur  $[BC]$  tel que  $BH$  corresponde à la plus petite des valeurs entre  $x$  et 1, enfin la perpendiculaire à  $(BC)$  menée par  $H$ , qui rencontre le cercle en un point  $A$ . D'après le théorème de Thalès sur le cercle<sup>26</sup>, le triangle  $ABC$  est rectangle en  $A$ .

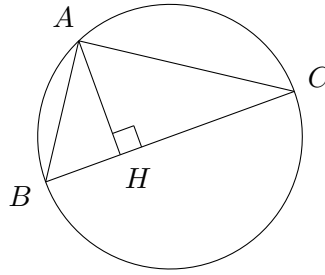
Les triangles  $ABC$  et  $ABH$  sont donc semblables parce qu'ils sont rectangles avec un angle (autre que l'angle droit) égal. Nous avons donc proportionnalité des longueurs des côtés :

$$\frac{AB}{BH} = \frac{BC}{AB}, \quad (19.265)$$

ce qui donne  $AB^2 = BC \times BH = x$  ( $BC$  et  $BH$  valent respectivement 1 et  $x$  ou le contraire).

25. Théorème 12.169.

26. Théorème 12.170.



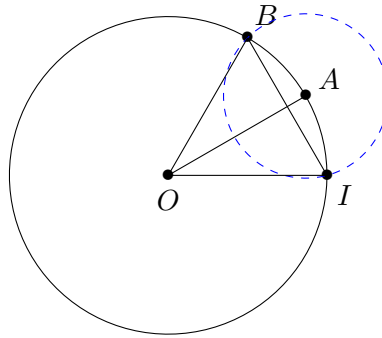
△

ExAHCooELGGPa

**Exemple 19.82** (Duplication d'un angle).

Si un angle  $\alpha$  est constructible, nous allons construire les angles  $2\alpha$ ,  $3\alpha$ , etc. Pour cela nous considérons un cercle de centre  $O$  et les points  $A$  et  $I$  sur le cercle tels que  $\widehat{AOI} = \alpha$ . Le cercle de centre  $A$  et de rayon  $AI$  intersecte le cercle de départ en les points  $I$  et  $B$ .

Le point  $A$  est à égale distance de  $B$  et  $I$ ; le point  $O$  également. Donc la droite  $(OA)$  est médiatrice du segment  $[BI]$ . Par conséquent elle est la hauteur du triangle isocèle  $OBI$ . L'angle  $\widehat{BOA}$  est alors le même que  $\widehat{AOI}$ ; par conséquent  $\widehat{BOI} = 2\alpha$ .



En traçant le cercle de centre  $B$  et de rayon  $BA$ , nous continuons et nous construisons  $3\alpha$ . △

L'exemple suivant qui permet d'additionner des angles repose sur le fait que deux cordes de mêmes longueurs sous-tendent des angles égaux, et est une adaptation simple de la duplication d'angle.

ExOVDooXnWPDl

**Exemple 19.83** (Addition d'angles).

Quitte à soustraire ou additionner un certain nombre de fois  $90^\circ$ , nous supposons que les deux angles donnés sont entre  $0^\circ$  et  $90^\circ$ .

Soient  $A, B, I$  sur un cercle de centre  $O$ , et nous notons  $\alpha = \widehat{OAI}$ ,  $\beta = \widehat{BOI}$ . Nous traçons le cercle de centre  $B$  et de rayon  $AI$ ; il intersecte le cercle en des points  $K_1$  et  $K_2$ . Les angles  $\widehat{K_1OB}$  et  $\widehat{K_2OB}$  sont tous deux égaux à  $\alpha$ .

Les angles  $\widehat{K_1OA}$  et  $\widehat{K_2OA}$  sont égaux à  $\beta - \alpha$  et  $\beta + \alpha$ . △

**19.84.**

Notez que l'exemple 19.83 donne un moyen de construire  $\alpha + \beta$  et  $\alpha - \beta$  ensemble. Il ne permet pas de savoir lequel est  $\alpha + \beta$  et lequel est  $\alpha - \beta$ .

## 19.5.2 Nombres constructibles

ThoRHFooZsLbqd

**Théorème 19.85** (Wantzel[642]).

Le réel  $a$  est constructible si et seulement si il existe une suite finie de corps  $\mathbb{L}_i$  tels que

- (1)  $\mathbb{L}_0 = \mathbb{Q}$ ,
- (2)  $\mathbb{L}_{i+1}$  est une extension quadratique<sup>27</sup> de  $\mathbb{L}_i$

27. C'est-à-dire une extension finie de degré 2.

(3)  $a \in \mathbb{L}_n$ .

*Démonstration.* Soit  $\mathbb{K}$  un corps de nombres constructibles (par exemple  $\mathbb{Q}$ ) ; nous notons  $E_{\mathbb{K}}$  l'ensemble des points de  $\mathbb{R}^2$  dont les coordonnées sont dans  $\mathbb{K}$ . Ce sont des points forcément constructibles.

- (1) **Intersection de droites** Si  $A, B \in E_{\mathbb{K}}$  alors la droite  $(AB)$  a pour équation  $ax + by + c = 0$  avec  $a, b, c \in \mathbb{K}$ , et le point d'intersection entre deux droites est donné par la solution du système

$$\begin{cases} ax + by + c = 0 \\ a'x + b'y + c' = 0, \end{cases} \quad \begin{matrix} (19.266a) \\ (19.266b) \end{matrix}$$

dont les solutions sont encore dans  $\mathbb{K}$ .

- (2) **Intersection droite-cercle** L'équation d'un cercle est de la forme

$$(x - u)^2 + (y - v)^2 = r^2 \quad (19.267)$$

où  $(u, v) \in E_{\mathbb{K}}$  et  $r$  est la distance entre deux points de  $E_{\mathbb{K}}$  ; donc  $r^2 \in \mathbb{K}$ . En développant et en redéfinissant  $u, v$  nous voyons que tous les cercles à considérer ont une équation de la forme

$$x^2 + y^2 + ux + vy + t = 0 \quad (19.268)$$

avec  $u, v, t \in \mathbb{K}$ . Il s'agit de voir où sont les solutions du système

$$\begin{cases} ax + by + c = 0 \\ x^2 + y^2 + ux + vy + t = 0. \end{cases} \quad \begin{matrix} (19.269a) \\ (19.269b) \end{matrix}$$

Si  $a \neq 0$  alors nous pouvons opérer la substitution  $x = -(c + by)/a$  et obtenir l'équation suivante pour  $y$  :

$$\left(\frac{-c - by}{a}\right)^2 + y^2 + u\left(\frac{-c - by}{a}\right) + vy + t = 0. \quad (19.270)$$

Cette équation est de la forme  $P(y) = 0$  où  $P$  est un polynôme du second degré à coefficients dans  $\mathbb{K}$ .

Si  $a = 0$  nous substituons  $y$  au lieu de  $x$  et le résultat est le même.

C'est le moment de relire la proposition 6.103 qui nous assure que si le réel  $\alpha$  est une solution de  $P(y) = 0$  hors de  $\mathbb{K}$ , alors l'extension  $\mathbb{K}(\alpha)$  est de degré 2 parce que le polynôme minimal de  $\alpha$  est de degré 2 :

$$[\mathbb{K}[\alpha] : \mathbb{K}] = 2. \quad (19.271)$$

De plus  $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$ .

- (3) **Intersection cercle-cercle** Le système

$$\begin{cases} x^2 + y^2 + ux + vy + t = 0 \\ x^2 + y^2 + ax + by + c = 0 \end{cases} \quad \begin{matrix} (19.272a) \\ (19.272b) \end{matrix}$$

est équivalent au système (en substituant la seconde équation par la différence entre les deux)

$$\begin{cases} x^2 + y^2 + ux + vy + t = 0 \\ (u - a)x + (v - b)y + t - c = 0 \end{cases} \quad \begin{matrix} (19.273a) \\ (19.273b) \end{matrix}$$

qui est à nouveau une intersection entre un cercle et une droite.

Passons à la conclusion. Si  $\alpha$  est un nombre constructible, alors il apparaît dans les coordonnées d'un point de  $C_m(E_{\mathbb{Q}})$  pour un certain  $m$ . Nous supposons (pour la récurrence) que chaque réel constructible en  $m - 1$  étapes possède sa pile d'extension quadratiques en partant de  $\mathbb{Q}$ . Le nombre  $\alpha$  vérifie donc

$$a\alpha^2 + b\alpha + c = 0 \quad (19.274)$$

avec  $a, b, c$  trois réels constructibles en  $m - 1$  étapes. L'hypothèse de récurrence donne donc des piles d'extensions

$$\begin{array}{lll} \mathbb{L}_0 = \mathbb{Q} & \mathbb{L}'_0 = \mathbb{Q} & \mathbb{L}''_0 = \mathbb{Q} \\ L_{i+1} = \mathbb{L}_i(a_i) & \mathbb{L}'_{i+1} = \mathbb{L}'_i(b_i) & \mathbb{L}''_{i+1} = \mathbb{L}''_i(c_i) \\ a \in \mathbb{L}_n & b \in \mathbb{L}'_{n'} & c \in \mathbb{L}''_{n''}. \end{array} \quad (19.275)$$

Nous considérons donc la suite d'extensions de  $\mathbb{Q}$  qui consiste à étendre successivement par les nombres  $a_1, \dots, a_n, b_1, \dots, b_{n'}, c_1, \dots, c_{n''}$  tout en excluant les doublons : il est possible que par exemple  $b_3$  soit déjà dans  $\mathbb{Q}(a_1, \dots, a_n)$ . Cela nous fournit une suite d'extensions

$$\begin{array}{l} \mathbb{M}_0 = \mathbb{Q} \\ \mathbb{M}_{i+1} = \mathbb{M}_i(\alpha_i) \\ a, b, c \in \mathbb{M}_n. \end{array} \quad (19.276)$$

Ici le  $n$  n'est pas spécialement le même que celui plus haut. Maintenant,  $\alpha$  est dans une extension quadratique de  $\mathbb{M}_n$ .

Pour la réciproque, nous supposons avoir une tour d'extensions quadratiques  $\mathbb{L}_0 = \mathbb{Q}$ ,  $\mathbb{L}_i$ , et  $\alpha \in \mathbb{L}_n$ , et nous voulons prouver que  $\alpha$  est constructible. Nous y allons par récurrence : si  $n = 0$  alors  $\alpha$  est rationnel, et il est constructible par la proposition 19.79.

Supposons que tous les points à coordonnées dans  $\mathbb{L}_i$  sont constructibles. Alors, nous allons prouver que les éléments de  $\mathbb{L}_{i+1}$  le sont également. Puisque  $\alpha$  est solution d'une équation de degré 2 à coefficients dans  $\mathbb{L}_i$ , nous avons

$$a\alpha^2 + b\alpha + c = 0 \quad (19.277)$$

pour certains  $a, b, c \in \mathbb{L}_i$  et donc

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \quad (19.278)$$

Les exemples 19.80 et 19.81 montrent que les produits et les racines carrées de nombres constructibles sont constructibles. Donc  $\alpha$  est constructible.  $\square$

### 19.5.3 Polygones constructibles

#### Définition 19.86.

Un angle  $\alpha$  est **constructible** si le nombre  $\cos(\alpha)$  est constructible.

La raison est qu'il suffit de prendre la perpendiculaire à l'axe horizontal.

LemMAHooXcOCpr

#### Lemme 19.87 ([126]).

Soient  $m$  et  $n$ , deux nombres premiers entre eux. L'angle  $\frac{2\pi}{mn}$  est constructible si et seulement si les angles  $\frac{2\pi}{m}$  et  $\frac{2\pi}{n}$  sont constructibles.

*Démonstration.* (1) **Sens direct** Il suffit de pouvoir multiplier un angle par un entier, ce qui est fait dans l'exemple 19.82.

(2) **Sens réciproque** Le théorème de Bézout 1.349 nous donne  $a, b \in \mathbb{Z}$  tels que  $an + bm = 1$ . Cela donne immédiatement

$$\frac{1}{mn} = \frac{a}{m} + \frac{b}{n}, \quad (19.279)$$

et donc

$$\frac{2\pi}{mn} = a \frac{2\pi}{m} + b \frac{2\pi}{n}, \quad (19.280)$$

ce qui fait que l'angle  $2\pi/mn$  est une combinaison entière d'angles constructibles. Il est donc constructible par l'exemple 19.83.  $\square$

LemUKNooSBzDyY

**Lemme 19.88.**

Soit un entier  $n \geq 3$  ayant

$$n = \prod_{i=1}^k p_i^{\alpha_i} \quad (19.281)$$

comme décomposition en facteurs premiers. Le polynôme régulier à  $n$  côtés est constructible si et seulement si les angles  $\frac{2\pi}{p_i^{\alpha_i}}$  sont constructibles.

*Démonstration.* Le polygone régulier à  $n$  côtés est constructible si et seulement si l'angle  $\frac{2\pi}{n}$  est constructible, c'est-à-dire si l'angle

$$\frac{2\pi}{\prod_{i=1}^k p_i^{\alpha_i}} \quad (19.282)$$

est constructible. Par récurrence sur le lemme 19.87, cet angle est constructible si et seulement si les angles  $\frac{2\pi}{p_i^{\alpha_i}}$  sont constructibles.  $\square$

ThoTWAooEsLjJu

**Théorème 19.89** (Gauss-Wantzel[126]).

Soit  $\alpha \in \mathbb{N} \setminus \{0\}$ .

ItemFSEooONDFrSi

(1) L'angle  $\frac{2\pi}{2^\alpha}$  est constructible.

ItemFSEooONDFrSii

(2) Si  $p$  est premier et  $p \neq 2$ , alors l'angle  $\frac{2\pi}{p^\alpha}$  est constructible si et seulement si  $\alpha = 1$  et  $p$  est un **nombre de Fermat**, c'est-à-dire de la forme  $1 + 2^{(2^\beta)}$  pour  $\beta \in \mathbb{N}$ .

ItemFSEooONDFrSiii

(3) Le polygone régulier à  $n$  côtés est constructible si et seulement si  $n$  est le produit d'une puissance de 2 et d'un nombre fini de nombres premiers de Fermat distincts.

*Démonstration.* Le point (1) est une construction de bissectrice, et le point (3) consistera à remettre en place différents morceaux. Le gros de la preuve est donc consacré à (2).

(1) **Sens direct** Nous supposons que l'angle  $\frac{2\pi}{p^\alpha}$  est constructible ; alors le nombre  $\cos(2\pi/p^\alpha)$  l'est également et le théorème de Wantzel 19.85 nous indique que

$$[\mathbb{Q}(\cos(2\pi/p^\alpha)) : \mathbb{Q}] = 2^m \quad (19.283)$$

pour un certain  $m$ . Posons  $q = p^\alpha$  et  $\omega = e^{2i\pi/q}$ . Grâce au corolaire 19.26, nous savons que le polynôme minimal de  $\omega$  est le polynôme cyclotomique  $\phi_q$  dont le degré est <sup>28</sup>

$$\varphi(q) = \varphi(p^\alpha) = p^{\alpha-1}(p-1). \quad (19.284)$$

Par conséquent

$$[\mathbb{Q}(\omega) : \mathbb{Q}] = p^{\alpha-1}(p-1). \quad (19.285)$$

Mais par ailleurs

$$\omega + \omega^{-1} = 2\cos(2\pi/q), \quad (19.286)$$

EqJDBooHURUQa

donc  $\cos(2\pi/q) \in \mathbb{Q}(\omega)$ . Et en multipliant (19.286) par  $\omega$  nous trouvons le polynôme annulateur suivant pour  $\omega$  :

$$\omega^2 - 2\cos\left(\frac{2\pi}{q}\right)\omega + 1 = 0. \quad (19.287)$$

Cela signifie que  $\omega$  est de degré 2 dans  $\mathbb{Q}(\cos(2\pi/q))$ , c'est-à-dire

$$[\mathbb{Q}(\omega) : \mathbb{Q}(\cos(2\pi/q))] = 2. \quad (19.288)$$

EqSJXooRCLJyt

28. Voir juste en dessous de la définition 19.19.

En remettant bout à bout et en utilisant la propriété multiplicative des degrés des extensions<sup>29</sup>,

$$\underbrace{[\mathbb{Q}(\omega) : \mathbb{Q}]}_{p^{\alpha-1}(p-1)} = \underbrace{[\mathbb{Q}(\omega) : \mathbb{Q}(\cos(2\pi/p^\alpha))]}_2 \underbrace{[\mathbb{Q}(\cos(2\pi/p^\alpha)) : \mathbb{Q}]}_{2^m}, \quad (19.289)$$

donc  $p^{\alpha-1}(p-1) = 2^{m+1}$ , mais comme  $p$  est premier et impair,  $\alpha = 1$  et  $p = 2^{m+1} + 1$ . Par ailleurs  $m+1$  est un entier et nous nous proposons de mettre en facteur la puissance de 2 dans son développement en facteurs premiers :  $m+1 = \lambda 2^\beta$  avec  $\beta \in \mathbb{N}$  et un certain nombre impair  $\lambda \in \mathbb{N}^*$ . Nous avons :

$$p = 2^{\lambda 2^\beta} + 1 = \left(2^{2^\beta}\right)^\lambda + 1, \quad (19.290)$$

mais le lemme 3.124(1) nous indique alors que  $1 + 2^{2^\beta}$  divise  $1 + \left(2^{2^\beta}\right)^\lambda = p$ . Mais vu que  $p$  est premier, il ne peut être divisé que par  $p$  lui-même et donc  $\lambda$  soit être égal à 1 et nous avons

$$p = 1 + 2^{2^\beta}, \quad (19.291)$$

ce qui signifie que  $p$  est un nombre de Fermat premier.

- (2) **Sens réciproque** Nous supposons que  $p$  est un nombre premier de la forme  $p = 1 + 2^{2^\beta}$  et nous devons prouver que l'angle  $\frac{2\pi}{p}$  est constructible. Pour cela nous posons tout de suite  $n = 2^\beta$  et  $\omega = e^{2i\pi/p}$ . Comme dans la première partie nous nous souvenons que le polynôme minimal de  $\omega$  est le polynôme cyclotomique  $\phi_p$  de degré  $p-1$  ; donc

$$[\mathbb{Q}(\omega) : \mathbb{Q}] = p-1. \quad (19.292)$$

- (2a) **Un groupe d'automorphismes** Nous considérons le groupe  $G = \text{Aut}_{\mathbb{Q}}(\mathbb{Q}(\omega))$  des automorphismes du corps  $\mathbb{Q}(\omega)$  agissant sur  $\mathbb{Q}$  comme l'identité. Tous les éléments de  $\mathbb{Q}(\omega)$  étant des polynômes en  $\omega$  (proposition 6.103), un élément  $g \in G$  est uniquement déterminé par  $g(\omega)$ , et de plus  $g(0) = 0$  ainsi que  $\phi_p(\omega) = 0$  et que  $\phi_p$  commute avec  $g$ , donc

$$\phi_p(g(\omega)) = g(\phi_p(\omega)) = g(0) = 0, \quad (19.293)$$

ce qui signifie que  $g(\omega)$  est une racine du polynôme cyclotomique  $\phi_p$ . Par définition 19.19, les racines sont  $\{\omega, \omega^2, \dots, \omega^{p-1}\}$ , ce qui signifie que l'action de  $g$  consiste à élever  $\omega$  à une certaine puissance entre 1 et  $p-1$ . Le groupe  $G$  est donc d'ordre  $|G| = p-1$  et a pour éléments

$$g_k : \omega \mapsto \omega^k \quad (19.294)$$

avec  $k = 1, \dots, p-1$ . L'application<sup>30</sup>

$$\begin{aligned} \psi : G &\rightarrow (\mathbb{Z}/p\mathbb{Z})^* \\ g_k &\mapsto [k]_p. \end{aligned} \quad (19.295)$$

est un morphisme surjectif entre deux groupes finis de même cardinal, donc c'est un isomorphisme. Le corolaire 19.34 nous donne de plus l'isomorphisme  $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$ , ce qui fait que  $G$  a un élément d'ordre  $p-1$  (parce que  $\mathbb{Z}/(p-1)\mathbb{Z}$  en a un). Nous notons  $g_0$  cet élément.

- (2b) **La tour d'extensions** À partir de cet élément  $g_0 \in G$  nous définissons avec  $0 \leq i \leq n$  :

$$\mathbb{K}_i = \{z \in \mathbb{Q}(\omega) \text{ tel que } g_0^{2^i}(z) = z\}. \quad (19.296)$$

Ces ensembles sont bien des corps parce que  $g_0$  est un morphisme :  $g_0(zz') = g_0(z)g_0(z')$  ; on en déduit immédiatement que  $g_0^2(zz') = g_0^2(z)g_0^2(z')$ .

29. Proposition 6.63.

30. Pour rappel, la notation  $[k]_p$  est la classe de l'entier  $k$  modulo  $p$ , qui est un élément de  $\mathbb{Z}/p\mathbb{Z}$ .

- (2c)  $\mathbb{K}_n = \mathbb{Q}(\omega)$  Cela est dû au fait que  $g_0$  est, par définition, d'ordre  $2^n$ , ce qui signifie que  $g_0^{2^n}(z) = z$  pour tout  $z \in \mathbb{Q}(\omega)$ .
- (2d)  $\mathbb{K}_0 = \mathbb{Q}$  Puisque les éléments de  $G$  laissent  $\mathbb{Q}$  invariant nous avons forcément  $\mathbb{Q} \subset \mathbb{K}_0$ . Nous nous attelons maintenant à prouver l'inclusion inverse. Nous savons par la proposition 6.103(2) que  $\{\omega^i\}_{i=0,\dots,p-2}$  est une base<sup>31</sup> de  $\mathbb{Q}(\omega)$ . Vu que  $g_0: \mathbb{Q}(\omega) \rightarrow \mathbb{Q}(\omega)$  est un isomorphisme, l'image d'une base est une base :

$$\{g_0(\omega)^i\}_{i=0,\dots,p-2} \quad (19.297)$$

est également une base de  $\mathbb{Q}(\omega)$ . Soit  $z \in \mathbb{K}_0$  et décomposons-le dans cette base<sup>32</sup> :

$$z = \sum_{i=0}^{p-2} \lambda_i g_0(\omega)^i; \quad (19.298) \quad \text{EqTMZqoCUySZb}$$

nous appliquons  $g_0$  à cette égalité en tenant compte du fait que  $g_0(z) = z$  :

$$z = \sum_{i=0}^{p-2} \lambda_i g_0(\omega)^{i+1}. \quad (19.299) \quad \text{EqGGDooMAngMs}$$

En identifiant les coefficients (et en remarquant que le dernier terme dans (19.298) est le premier dans (19.299)) nous voyons que tous les coefficients sont égaux :

$$z = \lambda_0 g_0(\omega + \dots + \omega^{p-1}) = \lambda_0 g_0(\phi_p(\omega) - 1) = -\lambda_0 \in \mathbb{Q}. \quad (19.300)$$

Dans cette chaîne d'égalités nous avons utilisé le fait que  $\phi_p(X) = 1 + X + \dots + X^{p-1}$  (corolaire 19.24) et que  $\phi_p(\omega) = 0$  par définition.

- (2e)  $\mathbb{K}_i \subset \mathbb{K}_{i+1}$  **strictement** Nous avons une inclusion pour la simple raison que si  $z \in \mathbb{K}_i$ , alors  $g_0^{2^i}(z) = z$ . Par conséquent :

$$g_0^{2^{i+1}}(z) = g_0^{2 \times 2^i}(z) = (g_0^{2^i})^2(z) = z. \quad (19.301)$$

Afin de voir que l'inclusion est stricte nous montrons que l'élément

$$x = \sum_{k=0}^{2^{n-i}-1} g_0^{k2^{i+1}}(\omega) \quad (19.302) \quad \text{EqTBLooWpeGgk}$$

est dans  $\mathbb{K}_{i+1} \setminus \mathbb{K}_i$ . Dans la base  $\{1, \omega, \dots, \omega^{p-1}\}$  l'élément  $x$  a une composante  $\omega$  avec le terme  $k = 0$ , mais si on lui applique  $g_0^{2^i}$  nous obtenons

$$g_0^{2^i}(x) = \sum_{k=0}^{2^{n-i-1}-1} g_0^{2^i(1+2k)}(\omega). \quad (19.303)$$

Une composante  $\omega$  pour cela demanderait d'avoir  $2^i(1+2k) = \lambda 2^n$  avec  $\lambda \in \mathbb{N}$ ; cela demanderait

$$k = \lambda 2^{n-i-1} - \frac{1}{2} \quad (19.304)$$

ce qui est impossible. Donc  $x \notin \mathbb{K}_i$ .

Nous montrons que  $x \in \mathbb{K}_{i+1}$  de la même manière :

$$g_0^{2^{i+1}}(x) = \sum_{k=0}^{2^{n-i-1}-1} g_0^{2^{i+1}(k+1)}(\omega). \quad (19.305) \quad \text{eqXHSoGyrfNul}$$

Soit  $k_0$  entre 0 et  $2^{n-i-1} - 1$ ; nous voulons trouver un  $k$  entre 0 et  $2^{n-i-1} - 1$  tel que  $k_0 2^{i+1} = 2^{i+1}(k+1)$ ; cela est très simple : il suffit de prendre  $k = k_0 - 1$  tant que  $k_0 \neq 0$ . Si  $k_0 = 0$  alors cela correspond au terme  $\omega$  dans (19.302), et il se trouve dans (19.305) avec  $k = 2^{n-i-1} - 1$ . Donc  $g_0^{2^{i+1}}(x) = x$  et  $x \in \mathbb{K}_{i+1}$ .

31. Ici [126] parle de  $\mathbb{Q}(\omega)/\mathbb{Q}$ . Dans ce cas il me semble qu'il faille faire partir les valeurs de  $i$  de 0 et non de 1.

32. Il s'agit d'une base en tant qu'espace vectoriel sur  $\mathbb{Q}$ , donc les  $\lambda_i$  sont dans  $\mathbb{Q}$ .

- (2f) **Les degrés dans la tour** Par définition de  $n$  et de  $\omega$ , et par la propriété multiplicative des degrés<sup>33</sup> nous avons :

$$2^n = p - 1 = [\mathbb{Q}(\omega) : \mathbb{Q}] = \underbrace{[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}] \dots [\mathbb{K}_1 : \mathbb{Q}]}_{n \text{ facteurs}}. \quad (19.306)$$

Un produit de  $n$  facteurs entiers tous strictement plus grands que zéro doit valoir  $2^n$ . Ils doivent donc tous valoir 2 et nous avons en particulier

$$[\mathbb{K}_{n-1} : \mathbb{Q}] = 2^{n-1}. \quad (19.307)$$

- (2g)  $\cos(2\pi/p) \in \mathbb{K}_{n-1}$  Soit  $f = g_0^{2^n-1}$  ; puisque l'action de  $g_0$  (comme de tous les éléments de  $G$ ) est de décaler les puissances de  $\omega$ , il existe  $\lambda \in \mathbb{N}$  tel que  $f(\omega) = \omega^\lambda$ . Mais  $f^2 = \text{Id}$  donc

$$\omega = f^2(\omega) = \omega^{2\lambda}, \quad (19.308)$$

ce qui donne  $1 = \omega^{\lambda^2-1}$  ou encore que  $p$  divise  $\lambda^2 - 1$  parce que  $\omega$  est une racine primitive  $p$ -ième de l'unité. Par conséquent  $[\lambda^2 - 1]_p = 0$  ce qui donne  $[\lambda]_p = \pm[1]_p$ , mais comme  $f \neq \text{Id}$  nous avons

$$[\lambda]_p = -[1]_p. \quad (19.309)$$

Ce que donne  $f(\omega) = \omega^{-1}$ .

Par ailleurs  $\mathbb{Q}(\omega)$  est un corps, donc il contient

$$\cos\left(\frac{2\pi}{p}\right) = \frac{1}{2}(\omega + \omega^{-1}). \quad (19.310)$$

Nous en déduisons que  $\cos(2\pi/p)$  est un point fixe de  $f = g_0^{2^n-1}$  :

$$f(\cos(2\pi/p)) = \frac{1}{2}(f(\omega) + f(\omega^{-1})) = \cos(2\pi/p). \quad (19.311)$$

Être un point fixe de  $g_0^{2^n-1}$  signifie être un élément de  $\mathbb{K}_{n-1}$  :

$$\cos\left(\frac{2\pi}{p}\right) \in \mathbb{K}_{n-1}. \quad (19.312)$$

- (2h) **Questions de degrés** Nous avons alors les (non)inclusions suivantes :

$$\mathbb{Q}(\cos(2\pi/p)) \subset \mathbb{K}_{n-1} \subsetneq \mathbb{K}_n = \mathbb{Q}(\omega), \quad (19.313)$$

ce qui fait que

$$1 < [\mathbb{Q}(\omega) : \mathbb{K}_{n-1}] \leq [\mathbb{Q}(\omega) : \mathbb{Q}(\cos(2\pi/p))] = 2 \quad (19.314) \quad \text{EqGEKooJBAQbg}$$

La première inégalité est le fait que  $\mathbb{Q}(\omega)$  n'est pas égal à  $\mathbb{K}_{n-1}$ . La dernière égalité se démontre de la même manière que (19.288) (ici  $\alpha = 1$ ). Les inégalités de (19.314) sont donc en réalité des égalités :

$$[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}] = [\mathbb{Q}(\omega) : \mathbb{Q}(\cos(2\pi/p))] = 2. \quad (19.315)$$

Cela, combiné au fait que  $\mathbb{Q}(\cos(2\pi/p)) \subset \mathbb{K}_{n-1}$  donne

$$\mathbb{K}_{n-1} = \mathbb{Q}(\cos(2\pi/p)). \quad (19.316)$$

Mais nous savons déjà que  $[\mathbb{K}_{n-1} : \mathbb{Q}] = 2^{n-1}$ , donc

$$[\mathbb{Q}(\cos(2\pi/p)) : \mathbb{Q}] = 2^{n-1} \quad (19.317)$$

et le nombre  $\cos(2\pi/p)$  est bien sûr le sommet d'une tour d'extensions quadratiques partant de  $\mathbb{Q}$ . Il est donc constructible par le théorème de Wantzel 19.85.

33. Proposition 6.63



Il est maintenant l'heure de conclure en prouvant le point (3). D'abord si  $n$  est un produit de nombres premiers de Fermat distincts, alors  $n = p_1 \dots p_m$  et l'angle  $\frac{2\pi}{n}$  est constructible si et seulement si chacun des angles  $\frac{2\pi}{p_k}$  est constructible (lemme 19.88), ce qui est le cas d'après la partie (2). Si  $n$  est le produit d'une puissance de 2 avec un produit de nombres premiers de Fermat, le polygone à  $n$  côtés est tout autant constructible : il suffit de bissecter les angles.

À l'inverse si le polynôme à  $n$  côtés (avec  $n$  impair) est constructible, alors  $n = \prod_i p_i^{\alpha_i}$ , dont l'angle est constructible si et seulement si  $\alpha_i = 1$  et  $p_i$  est un nombre premier de Fermat. Et si le polygone à  $n$  côtés ( $n$  pair) est constructible, alors  $n = 2^\lambda \prod_i p_i^{\alpha_i}$ . Ce polygone est constructible si et seulement si le polygone à  $\prod_i p_i^{\alpha_i}$  côtés est constructible : il suffit de regrouper les côtés par deux.  $\square$

**Remarque 19.90.**

D'après Wikipédia[643], les seuls nombres de Fermat connus pour être premiers sont

$$F_0 = 3 \quad (19.318a)$$

$$F_1 = 5 \quad (19.318b)$$

$$F_2 = 17 \quad (19.318c)$$

$$F_3 = 257 \quad (19.318d)$$

$$F_4 = 65537. \quad (19.318e)$$

Pour les autres, essentiellement on ne sait pas. Il n'est même pas sûr qu'il y en ait d'autres. Le problème des polygones constructibles n'est donc pas encore tout à fait terminé.

**Remarque 19.91.**

Les angles  $\frac{2\pi}{p^\alpha}$  avec  $p$  premier et  $\alpha > 1$  ne sont pas constructibles. Il n'est donc pas possible de trisecter l'angle  $\frac{2\pi}{3}$ . Voilà qui règle un des vieux problèmes de l'antiquité.

# Chapitre 20

## Intégration sur des variétés

### 20.1 Variétés

#### 20.1.1 Introduction

Soit  $f : S^2 \rightarrow \mathbb{R}$  une fonction définie sur la sphère usuelle  $S^2 \subset \mathbb{R}^3$ . Une question naturelle est d'estimer la régularité de  $f$  ; est-elle continue, dérivable, différentiable ? Il n'existe pas de dérivée directionnelle étant donné que le quotient différentiel

$$\frac{f(x + \epsilon u_1, y + \epsilon u_2) - f(x, y)}{\epsilon}$$

n'a pas de sens pour un point  $(x + \epsilon u_1, y + \epsilon u_2)$  qui n'est pas –sauf valeurs particulières– dans la surface. Pour la même raison il n'est pas possible de parler de différentiabilité de cette manière. Comment faire, sans devoir étendre le domaine de définition de  $f$  à un voisinage de la sphère ? Une solution possible est de parler de la notion de variété.

Une variété est un objet qui ressemble, vu de près, à  $\mathbb{R}^m$  pour un certain  $m$ . En d'autres termes, on imagine une variété comme un recollement de morceaux de  $\mathbb{R}^m$  vivant dans un espace plus grand  $\mathbb{R}^n$ . Ces morceaux sont appelés des ouverts de carte, et l'application qui exprime la ressemblance à  $\mathbb{R}^m$  est l'application de carte.

#### 20.1.2 Définition, carte

DEFooJOAHooAxcLHe

##### Définition 20.1.

Une **variété** de dimension  $m$  de classe  $C^1$  est une partie  $M$  de  $\mathbb{R}^n$  ( $n \geq m$ ) munie d'un ensemble de paires  $\{U_\alpha, \varphi_\alpha\}$  où  $U_\alpha$  est un ouvert de  $\mathbb{R}^m$  et  $\varphi_\alpha : U_\alpha \rightarrow \mathbb{R}^n$  est une application vérifiant

- (1)  $\varphi_\alpha$  est une bijection entre  $U_\alpha$  et  $\varphi_\alpha(U_\alpha)$  ;
- (2) pour tout  $a \in M$ , il existe un  $\alpha$  tel que  $a \in \varphi_\alpha(U_\alpha)$  ;
- (3) pour tout  $\alpha, \beta$ , la partie  $\varphi_\alpha^{-1}(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta))$  est un ouvert de  $U_\alpha$  ;
- (4) L'application

$$\varphi_\beta^{-1} \circ \varphi_\alpha : \varphi_\alpha^{-1}(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)) \rightarrow U_\beta \quad (20.1)$$

est un  $C^1$ -difféomorphisme vers son image.

Notez que cette définition n'utilise pas du tout la structure de  $\mathbb{R}^n$  dans lequel se trouve  $M$ . Seulement la structure différentielle du  $\mathbb{R}^m$  depuis lequel les cartes partent. En particulier, nous n'avons pas dit si  $\varphi_\alpha : U_\alpha \rightarrow M$  était de classe  $C^1$  pour une topologie induite de  $\mathbb{R}^n$  vers  $M$ . Nous pouvons en réalité définir plus généralement une variété en remplaçant « une partie de  $\mathbb{R}^n$  » par « un ensemble » sans rien y changer.

La difficulté qui apparaît lorsque nous voulons dire que  $M$  est un ensemble quelconque au lieu d'une partie de  $\mathbb{R}^n$  est pour la mesure.

Nous verrons plus loin comment les cartes  $\varphi_\alpha$  peuvent transporter la mesure de  $U_\alpha$  (celle de Lebesgue dans  $\mathbb{R}^m$ ) vers une partie de  $\mathbb{R}^n$  en utilisant le déterminant de  $d\varphi_\alpha$ . Cela n'est possible

que parce que  $\mathbb{R}^n$  a une structure différentielle qui permet de définir  $d\varphi_\alpha$ . Si  $M$  est un ensemble quelconque, cette façon de faire n'est plus possible. Dans ce cas, il faut définir des fibrés tangents et cotangents puis en extraire une « forme volume ». Le choix de la forme volume n'étant pas canonique, il n'y a pas d'intégrale canonique sur une variété générale comme il y en a une sur une variété « dans  $\mathbb{R}^n$  ».

DEFooWABZooGIEDEV

### Définition 20.2.

Soit une variété  $M$  de dimension  $m$ . Une **carte** pour  $M$  est un couple  $(U, \varphi)$  où  $U$  est un ouvert de  $\mathbb{R}^m$  et  $\varphi: U \rightarrow M$  est une application

- (1) qui est une bijection entre  $U$  et son image,
- (2) pour tout  $\alpha$ ,

$$\varphi_\alpha^{-1} \circ \varphi: \varphi^{-1}(\varphi(U) \cap \varphi_\alpha(U_\alpha)) \rightarrow \varphi_\alpha^{-1}(\varphi(U) \cap \varphi_\alpha(U_\alpha)) \quad (20.2)$$

est un  $C^1$ -difféomorphisme.

### Proposition 20.3 ([1]).

Si  $(U_1, \varphi_1)$  et  $(U_2, \varphi_2)$  sont des cartes pour la variété  $M$ , alors en posant  $S = \varphi_1(U_1) \cap \varphi_2(U_2)$ , l'application

$$\varphi_2^{-1} \circ \varphi_1: \varphi_1^{-1}(S) \rightarrow \varphi_2^{-1}(S) \quad (20.3)$$

est un  $C^1$ -difféomorphisme.

*Démonstration.* Le fait que  $\varphi_2^{-1} \circ \varphi_1$  soit une bijection est dû au fait que nous ayons choisi les espaces de départ et d'arrivée pour que ce soit une bijection.

Soit  $a \in \varphi_1(U_1) \cap \varphi_2(U_2)$ . Soient  $\alpha$  tel que  $a \in \varphi_\alpha(U_\alpha)$  et un ouvert  $A$  autour de  $\varphi_\alpha^{-1}(a)$  tel que  $\varphi_\alpha(A) \subset \varphi_1(U_1) \cap \varphi_2(U_2)$ . Nous allons montrer que

$$\varphi_2^{-1} \circ \varphi_1: \varphi_1^{-1}(\varphi_\alpha(A)) \rightarrow \varphi_2^{-1}(\varphi_\alpha(A)) \quad (20.4)$$

est un  $C^1$ -difféomorphisme. Nous avons

$$\varphi_2^{-1} \circ \varphi_1 = \varphi_2^{-1} \circ \varphi_\alpha \circ \varphi_\alpha^{-1} \circ \varphi_1, \quad (20.5)$$

mais vu que  $\varphi_1$  et  $\varphi_2$  sont des cartes, les applications  $\varphi_2^{-1} \circ \varphi_\alpha$  et  $\varphi_\alpha \circ \varphi_1$  sont de classe  $C^1$  et d'inverses  $C^1$ . La composition l'est encore.  $\square$

Nous avons demandé qu'une variété n'admette que des cartes partant d'ouverts de  $\mathbb{R}^m$ . Aurions-nous pu admettre, dans la définition 20.2 que la carte parte d'un ouvert de  $\mathbb{R}^p$  avec  $p \neq m$ ? Non. Voici une résultat qui dit que si une carte part de  $\mathbb{R}^m$ , alors toutes les cartes doivent partir de  $\mathbb{R}^m$ .

### Proposition 20.4.

Soit une variété  $M$  de dimension  $m$  et une carte  $\varphi_\alpha: U_\alpha \rightarrow M$ . Si  $U$  est un ouvert de  $\mathbb{R}^p$  et si  $\varphi: U \rightarrow M$  est telle que  $\varphi_\alpha^{-1} \circ \varphi$  soit un difféomorphisme, alors  $p = m$ .

*Démonstration.* Soit  $A = \varphi_\alpha(U_\alpha) \cap \varphi(U)$ . Les parties  $\varphi_\alpha^{-1}(A)$  et  $\varphi^{-1}(U)$  sont des ouverts de  $\mathbb{R}^m$  et de  $\mathbb{R}^p$  respectivement. Par hypothèse, l'application

$$\varphi_\alpha^{-1} \circ \varphi: \varphi^{-1}(A) \rightarrow \varphi_\alpha^{-1}(A) \quad (20.6)$$

est un difféomorphisme entre l'ouvert  $\varphi^{-1}(A) \subset \mathbb{R}^p$  et l'ouvert  $\varphi_\alpha^{-1}(A) \subset \mathbb{R}^m$ . La proposition 10.312 implique que  $m = p$ .  $\square$

### 20.1.3 Ancienne définition

Cette section est à recycler.

**Définition 20.5.**

Soit  $\emptyset \neq M \subset \mathbb{R}^n$ ,  $1 \leq m < n$  et  $k \geq 1$ .  $M$  est une variété de classe  $C^k$  de dimension  $m$  si pour tout  $a \in M$ , il existe un voisinage ouvert  $U$  de  $a$  dans  $\mathbb{R}^n$ , et un ouvert  $V$  de  $\mathbb{R}^m$  tel que  $U \cap M$  soit le graphe d'une fonction  $f : V \subset \mathbb{R}^m \rightarrow \mathbb{R}^{n-m}$  de classe  $C^1$ , c'est-à-dire qu'il existe un réagencement des coordonnées  $(x_{i_1}, \dots, x_{i_m}, x_{i_{m+1}}, \dots, x_{i_n})$  avec

$$M \cap U = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \text{ tel que } (x_{i_1}, \dots, x_{i_m}) \in V \mid \begin{cases} x_{i_{m+1}} = f_1(x_{i_1}, \dots, x_{i_m}) \\ \vdots \\ x_{i_n} = f_{n-m}(x_{i_1}, \dots, x_{i_m}) \end{cases} \right\}$$

où  $V$  est un voisinage ouvert de  $(a_{i_1}, \dots, a_{i_m}) \in \mathbb{R}^m$ .

La littérature regorge de théorèmes qui proposent des conditions équivalentes à la définition d'une variété. Celle que nous allons le plus utiliser est la suivante

**Proposition 20.6.**

Soit  $M \subset \mathbb{R}^n$  et  $1 \leq m \leq n-1$ . L'ensemble  $M$  est une variété si et seulement si  $\forall a \in M$ , il existe un voisinage ouvert  $U$  de  $a$  dans  $\mathbb{R}^n$  et une application  $F : W \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$  où  $W$  est un ouvert tels que

- (1)  $F$  est un homéomorphisme de  $W$  vers  $M \cap U$ ,
- (2)  $F \in C^1(W, \mathbb{R}^n)$ ,
- (3) Le rang de  $dF(w) \in L(\mathbb{R}^m, \mathbb{R}^n)$  est de rang maximum (c'est-à-dire  $m$ ) en tout point  $w \in W$ .

Pour rappel, si  $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$  est une application linéaire, son rang<sup>1</sup> est la dimension de son image. Si  $A$  est la matrice d'une application linéaire, alors le rang de cette application linéaire est égal à la taille de la plus grande matrice carrée de déterminant non nul contenue dans  $A$ <sup>2</sup>.

La condition de rang maximum sert à éviter le genre de cas de la figure 20.1 qui représente l'image de l'ouvert  $[-1, 1]$  par l'application  $F(t) = (t^2, t^3)$ .

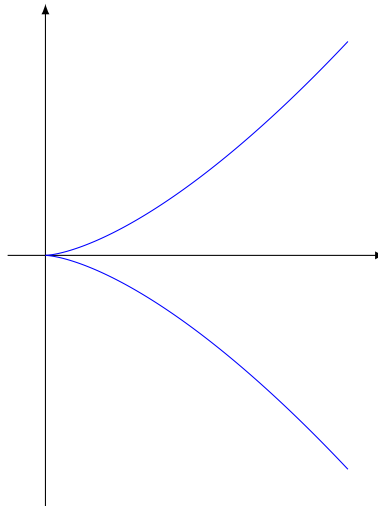


FIGURE 20.1: Quelque chose qui n'est pas de rang maximum et qui n'est pas une variété

La différentielle a pour matrice

$$dF(t) = (2t, 3t^2). \quad (20.7)$$

Le rang maximum est 1, mais en  $t = 0$ , la matrice vaut  $(0, 0)$  et son rang est zéro. Pour toute autre valeur de  $t$ , c'est bon.

Une autre caractérisation des variétés est donnée par la proposition suivante

---

1. Définition 4.48.  
2. Proposition 4.134

PropCarVarZerFonc

**Proposition 20.7.**

Soit  $M \subset \mathbb{R}^n$  et  $1 \leq m \leq n-1$ . L'ensemble  $M$  est une variété si et seulement si  $\forall a \in M$ , il existe un voisinage ouvert  $\mathcal{U}$  de  $a$  dans  $\mathbb{R}^n$  tel et une application  $G \in C^1(\mathcal{U}, \mathbb{R}^{n-m})$  tel que

- (1) le rang de  $dG_a \in L(\mathbb{R}^n, \mathbb{R}^{n-m})$  soit maximum (c'est-à-dire  $n-m$ ) en tout  $a \in M$ ,
- (2)  $M \cap \mathcal{U} = \{x \in \mathcal{U} \text{ tel que } G(x) = 0\}$ .

**20.1.4 Espace tangent**

DEFooMCEYooHuCADg

**Proposition-Définition 20.8** ([644]).

Soit une variété  $M$  de dimension  $m$  dans  $\mathbb{R}^n$ . Soit  $a \in M$ . L'ensemble

$$T_a M = \left\{ \gamma'(0) \text{ tel que } \begin{cases} \gamma \in C^1(]-1, 1[, M) \\ \gamma(0) = a \end{cases} \right\} \quad (20.8)$$

est un sous-espace vectoriel de dimension  $m$ , appelé **espace tangent** de  $M$  en  $a$ .

*Démonstration.* Pas de preuve pour le moment. Vous pouvez aider en m'envoyant une preuve. □

**20.2 Intégration****20.2.1 Le problème pour une intégration globale**

Soient une variété  $M$  et une partie  $A$  de  $M$ . Si nous avons une fonction  $f: A \rightarrow \mathbb{R}$ , nous voudrions définir une intégrale de  $f$  sur  $A$  en utilisant les cartes.

Il y a une petite complication. Supposons que  $A$  soit dans  $\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)$ . Pour la carte  $\varphi_\alpha$  nous aurions tendance à vouloir écrire

$$\int_A f = \int_{\varphi_\alpha^{-1}(A)} f \circ \varphi_\alpha \quad (20.9) \quad \text{EQooUEVHooOXgonw}$$

pour la mesure de Lebesgue induite de  $\mathbb{R}^m$  vers son ouvert  $U_\alpha$ . Pas de problème à ce que (20.9) soit bien définie. Le problème est que  $\alpha$  n'est pas à priori spécial par rapport à  $\beta$  et qu'il faudrait également que

$$\int_A f = \int_{\varphi_\beta^{-1}(A)} f \circ \varphi_\beta \quad (20.10) \quad \text{EQooEOJHooGLQHs}$$

Mais en utilisant le changement de variable (théorème 14.282(3)) pour le  $C^1$ -difféomorphisme  $\phi = \varphi_\alpha^{-1} \circ \varphi_\beta$ , nous avons

$$\int_{\varphi_\alpha^{-1}(A)} f \circ \varphi_\alpha = \int_{\varphi_\beta^{-1}(A)} (f \circ \varphi_\alpha \circ \varphi_\alpha^{-1} \circ \varphi_\beta) |J_\phi| = \int_{\varphi_\beta^{-1}(A)} (f \circ \varphi_\beta) |J_\phi|. \quad (20.11)$$

À moins d'une coïncidence extraordinaire sur les valeurs du jacobien, il n'y aura pas égalité entre (20.9) et (20.10).

Pour faire mieux, il faudra ajouter quelque chose qui compense l'arrivée du jacobien. C'est ce que nous allons faire maintenant.

**20.2.2 Intégrale sur une carte**

Dans un premier temps, nous allons définir l'intégrale sur une carte. Nous verrons plus tard comment combiner les cartes pour faire une intégrale sur une variété entière.

PROPooOAHWooAfxvyv

**Proposition-Définition 20.9** ([1]).

Soient deux cartes  $(U_1, \varphi_1)$  et  $(U_2, \varphi_2)$  pour la partie  $A$  d'une variété de dimension  $m$  dans  $\mathbb{R}^n$ . Soit une fonction mesurable  $f: A \rightarrow [0, +\infty]$ <sup>3</sup>.

3. Si  $f$  est seulement mesurable, allez voir les hypothèses du théorème 14.282(3).

Alors<sup>4</sup>

$$\int_{U_1} (f \circ \varphi_1)(x) \sqrt{\det((d\varphi_1)_x^*(d\varphi_1)_x)} dx = \int_{U_2} (f \circ \varphi_2)(y) \sqrt{\det((d\varphi_2)_y^*(d\varphi_2)_y)} dy \quad (20.12)$$

où les deux intégrales sont au sens de la mesure de Lebesgue sur des ouverts de  $\mathbb{R}^m$ .

Ce nombre est **l'intégrale** de  $f$  sur  $A$ , et est noté

$$\int_A f. \quad (20.13)$$

*Démonstration.* Vu que nous sommes en présence de cartes pour une variété, les applications  $\varphi_1: U_1 \rightarrow A$  et  $\varphi_2: U_2 \rightarrow A$  sont des bijections telles que  $\varphi_1^{-1} \circ \varphi_2: U_2 \rightarrow U_1$  soit un  $C^1$ -difféomorphisme.

Nous considérons le  $C^1$ -difféomorphisme  $\phi = \varphi_1^{-1} \circ \varphi_2$ . Notre but est d'utiliser le théorème de changement de variables 14.282(2).

Vu que  $\varphi_1$  est une bijection,  $(\varphi_1 \circ \varphi_1^{-1})(A) = A$  et nous avons

$$\phi^{-1}(\varphi_1^{-1}(A)) = \varphi_2^{-1}(A). \quad (20.14)$$

Ensuite,  $f \circ \varphi_1 \circ \phi = f \circ \varphi_2$ .

Jusqu'ici, rien de drôle me diriez-vous. Ok. Alors voyons un peu comment se passe le jacobien et ce qui se trouve sous la racine carrée.

L'application  $x \mapsto \det((d\varphi_1)_x^* \circ (d\varphi_1)_x)$  doit être composée avec  $\phi$  pour donner

$$y \mapsto \det((d\varphi_1)_{\phi(y)}^* \circ (d\varphi_1)_{\phi(y)}). \quad (20.15)$$

Attardons-nous un peu sur  $(d\varphi_1)_{\phi(y)}$ . Nous savons par la règle de différentiation en chaîne 10.304 que

$$d(\varphi_1 \circ \phi)_y = (d\varphi_1)_{\phi(y)} \circ (d\phi)_y. \quad (20.16)$$

Le lemme 10.311 nous dit que l'application linéaire  $d\phi_y: \mathbb{R}^m \rightarrow \mathbb{R}^m$  est inversible et que  $d\phi_y^{-1} = (d\phi^{-1})_{\phi(y)}$ . En composant les deux côtés de (20.16) par cela nous trouvons

$$(d\varphi_1)_{\phi(y)} = d(\varphi_1 \circ \phi)_y \circ (d\phi)_y^{-1} = (d\varphi_2)_y \circ (d\phi^{-1})_{\phi(y)}. \quad (20.17)$$

C'est le moment d'utiliser la proposition 8.39 à propos de la composition des applications adjointes. Ce qui est dans le déterminant est :

$$((d\varphi_2)_y \circ (d\phi^{-1})_{\phi(y)})^* \circ ((d\varphi_2)_y \circ (d\phi^{-1})_{\phi(y)}) = \underbrace{(d\phi^{-1})_{\phi(y)}^*}_{\mathbb{R}^m \rightarrow \mathbb{R}^m} \circ \underbrace{(d\varphi_2)_y^* \circ (d\varphi_2)_y}_{\mathbb{R}^m \rightarrow \mathbb{R}^m} \circ \underbrace{(d\phi^{-1})_{\phi(y)}}_{\mathbb{R}^m \rightarrow \mathbb{R}^m}. \quad (20.18)$$

Nous utilisons à présent un combo entre les propriétés des déterminants et de l'adjoint : les propositions 8.16(2) et 8.38 font sortir  $|\det((d\phi^{-1})_{\phi(y)})|$  de la racine carrée. Ne pas oublier la valeur absolue parce que  $\sqrt{x^2} = |x|$  et non  $\sqrt{x^2} = x$ . Cela pour dire que

$$\sqrt{\det((d\varphi_1)_{\phi(y)}^* \circ (d\varphi_1)_{\phi(y)})} = |\det((d\phi^{-1})_{\phi(y)})| \sqrt{\det((d\varphi_2)_y^* \circ (d\varphi_2)_y)}. \quad (20.19)$$

En ce qui concerne l'intégrale que nous voulons calculer, en mettant les bouts ensemble,

$$\begin{aligned} & \int_{\varphi_1^{-1}(A)} (f \circ \varphi_1)(x) \sqrt{\det(\dots)} dx \\ &= \int_{\varphi_2^{-1}(A)} (f \circ \varphi_2)(y) |\det((d\phi^{-1})_{\phi(y)})| \sqrt{\det((d\varphi_2)_y^* \circ (d\varphi_2)_y)} |\det(d\phi_y)| dy. \end{aligned} \quad (20.20)$$

Enfin, les déterminants se suppriment : dans la valeur absolue nous avons :

$$\det((d\phi^{-1})_{\phi(y)}) \det(d\phi_y) = \det((d\phi^{-1})_{\phi(y)} \circ d\phi_y) = \det(\text{Id}) = 1. \quad (20.21)$$

□

4. Voir la définition de l'adjoint 8.36.

**20.10.**

Notez que si  $A$  est un ouvert de  $\mathbb{R}^m$  lui-même, le nombre défini dans la définition 20.9 est l'intégrale usuelle. Pour le voir, choisir  $U = A$  ainsi que la carte

$$\begin{aligned}\varphi: U &\rightarrow A \\ x &\mapsto x.\end{aligned}\tag{20.22}$$

La différentielle de  $\varphi$  est l'identité, de telle sorte que  $\det((d\varphi)_x^*(d\varphi)_x) = 1$ . De même,  $f \circ \varphi = f$ . Il reste seulement  $\int_U f(x)dx$ .

Si nous voulons que la formule de la proposition 20.9 fournisse une définition raisonnable de  $\int_A f$ , il faut que si  $(U, \varphi)$  est une carte,

$$\det(d\varphi_x^* \circ d\varphi_x) \geq 0\tag{20.23}$$

pour tout  $x$ , et ne soit nul sur aucun ouvert de  $U$ .

**Lemme 20.11** ([1]).

Soit une carte  $(U, \varphi)$  d'une variété  $M$  de dimension  $m$ . Soit  $x \in U$ . Alors

- (1) L'opérateur  $d\varphi_x^* \circ d\varphi_x$  est autoadjoint semi-défini positif.
- (2) Il existe une base de  $\mathbb{R}^m$  formée de vecteurs propres de  $d\varphi_x^* \circ d\varphi_x$ . Les valeurs propres sont réelles et positives.
- (3) Pour tout  $x \in U$  nous avons

$$\det(d\varphi_x^* \circ d\varphi_x) > 0.\tag{20.24}$$

*Démonstration.* Point par point.

- (1) Soit  $x \in U$ . Nous notons  $A = d\varphi_x$ . Par la proposition 8.39 nous avons  $(AA^*)^* = AA^*$  du fait que  $(A^*)^* = A$ . Donc  $A^*A$  est autoadjoint. Sa matrice est donc symétrique par la proposition 8.36(2).

De plus, pour tout  $u \in \mathbb{R}^m$  nous avons

$$\langle A^*Au, u \rangle = \langle Au, Au \rangle \geq 0.\tag{20.25}$$

Cela implique que la matrice de  $A^*A$  est semi-définie positive par le lemme 8.233(2). L'opérateur  $A^*A$  est également semi-défini positif par la proposition 8.234.

- (2) L'existence de la base de vecteurs propres est le théorème 8.226. Ce théorème dit de plus que les valeurs propres sont réelles. Le fait qu'elles soient positives est le fait déjà prouvé que  $A^*A$  est semi-défini positif.
- (3) Dans la base de vecteurs propres, la matrice est diagonale avec des nombres positifs sur la diagonale. Le déterminant est alors le produit des valeurs propres. Voilà pourquoi le déterminant est positif.

Il nous reste à montrer que ce déterminant ne peut pas être nul. Si  $\det(d\varphi_a^* \circ d\varphi_a) = 0$ , alors l'application  $d\varphi_a^* \circ d\varphi_a$  a un noyau (proposition 8.16(2)), c'est-à-dire un vecteur  $u$  tel que  $d\varphi_a^* d\varphi_a(u) = 0$  en particulier,

$$0 = \langle d\varphi_a^* d\varphi_a u, u \rangle = \langle d\varphi_a u, d\varphi_a u \rangle = \|d\varphi_a u\|^2.\tag{20.26}$$

Donc  $d\varphi_a(u) = 0$ .

Mais  $\varphi: U \rightarrow \varphi(U)$  est bijective, donc l'application  $\varphi^{-1}: \varphi(U) \rightarrow U$  l'est aussi et vu que  $\varphi$  est une carte, l'application  $\varphi^{-1} \circ \varphi: U \rightarrow U$  est égale à l'identité. L'identité est une application linéaire dont la différentielle est elle-même, c'est-à-dire que

$$d(\varphi^{-1} \circ \varphi)_a = (d\varphi^{-1})_{\varphi(a)} \circ d\varphi_a = \text{Id}.\tag{20.27}$$

Si donc  $u \neq 0$ , nous devons avoir  $d\varphi_a(u) \neq 0$ .

□

### 20.2.3 Quelques expressions pour l'élément de volume

Ce que nous appelons « élément de volume » (pour des raisons qui apparaîtront plus tard) est le coefficient  $\sqrt{\det(d\varphi_x^* \circ d\varphi_x)}$ , introduit pour le besoin de compenser l'arrivée du jacobien en cas de changement de variables. Mais à quoi ressemble ce coefficient pour les variétés de petite dimension ?

C'est ce que nous allons voir maintenant.

#### 20.2.3.1 En dimension un

Nous supposons une variété de dimension un dans  $\mathbb{R}^n$ . Une carte est donc une application  $\varphi: \mathbb{R} \rightarrow \mathbb{R}^n$  dont nous notons  $\varphi_i$  les composantes. Nous avons, pour  $u \in \mathbb{R}$ ,

$$d\varphi_a(u) = \frac{d}{dt} [\varphi(a + tu)]_{t=0} = \frac{d}{dt} \begin{pmatrix} \varphi_1(a + tu) \\ \vdots \\ \varphi_n(a + tu) \end{pmatrix} = u \begin{pmatrix} \varphi'_1(a) \\ \vdots \\ \varphi'_n(a) \end{pmatrix} = u\varphi'(a). \quad (20.28)$$

Là dedans, vous vous souviendrez que  $a \in \mathbb{R}$  et que  $\varphi'(a) \in \mathbb{R}^n$ .

Nous devons savoir ce que vaut  $d\varphi_a^*x$  lorsque  $a \in \mathbb{R}$  et  $x \in \mathbb{R}^n$ . Pour tout  $u \in \mathbb{R}$  nous avons

$$\langle d\varphi_a x, u \rangle_{\mathbb{R}} = \langle x, d\varphi_a(u) \rangle_{\mathbb{R}^n} = u \langle x, \varphi'(a) \rangle_{\mathbb{R}^n} = u \varphi'(a) \cdot x = \langle \varphi'(a) \cdot x, u \rangle_{\mathbb{R}}. \quad (20.29)$$

Donc  $d\varphi_a^*(x) = \varphi'(a) \cdot x$ .

Question notation, nous avons noté  $\langle a, b \rangle_{\mathbb{R}} = ab$  et indifféremment  $\langle x, y \rangle_{\mathbb{R}^n} = x \cdot y$ .

L'application  $d\varphi_a^* \circ d\varphi_a$  est une application linéaire bijective entre  $\mathbb{R}$  et  $\mathbb{R}$ . Sa matrice est donc  $1 \times 1$  et elle se calcule en appliquant l'application au vecteur de base 1 de  $\mathbb{R}$  :

$$(d\varphi_a^* \circ d\varphi_a)(1) = d\varphi_a^*(\varphi'(a)) = \varphi'(a) \cdot \varphi'(a) = \|\varphi'(a)\|^2. \quad (20.30)$$

Nous avons donc

$$\det(d\varphi_a^* \circ d\varphi_a) = \|\varphi'(a)\|^2. \quad (20.31)$$

#### 20.2.3.2 En dimension quelconque

Soit une carte  $\varphi: \mathbb{R}^m \rightarrow \mathbb{R}^n$  pour une variété de dimension  $m$ . Les éléments de matrice de la différentielle d'une application sont donnés par la proposition 12.272 :

$$(d\varphi_a)_{ij} = \frac{\partial \varphi_i}{\partial x_j}(a). \quad (20.32)$$

Les éléments de la matrice de  $d\varphi_a^*$  sont ceux de la matrice transposée. En ce qui concerne la composition, c'est le produit des matrices :

$$(d\varphi_a^* \circ d\varphi_a)_{ij} = \sum_k (d\varphi_a^*)_{ik} (d\varphi_a)_{kj} = \sum_k (d\varphi_a)_{ki} (d\varphi_a)_{kj} = \sum_k \frac{\partial \varphi_k}{\partial x_i}(a) \frac{\partial \varphi_k}{\partial x_j}(a). \quad (20.33)$$

Cela pour écrire cette bonne formule :

$$(d\varphi_a^* \circ d\varphi_a)_{ij} = \frac{\partial \varphi}{\partial x_i}(a) \cdot \frac{\partial \varphi}{\partial x_j}(a). \quad (20.34)$$

#### 20.2.3.3 En dimension deux

Nous considérons maintenant une carte  $\varphi: U \rightarrow \mathbb{R}^3$  avec  $U \subset \mathbb{R}^2$ . En notant  $v_i = \partial_i \varphi(a)$  nous notons la formule (20.34) sous la forme

$$\det(d\varphi_a^* \circ d\varphi_a) = \det \begin{pmatrix} v_1 \cdot v_1 & v_1 \cdot v_2 \\ v_2 \cdot v_1 & v_2 \cdot v_2 \end{pmatrix} = \|v_1\|^2 \|v_2\|^2 - (v_1 \cdot v_2)^2. \quad (20.35)$$

L'identité de Lagrange de la proposition 10.43 nous donne alors

$$\sqrt{\det(d\varphi_a^* \circ d\varphi_a)} = \left\| \frac{\partial \varphi}{\partial x} \times \frac{\partial \varphi}{\partial y} \right\|. \quad (20.36)$$



### 20.2.3.4 En dimension trois

Nous considérons maintenant une carte  $\varphi: U \rightarrow \mathbb{R}^3$  avec  $U \subset \mathbb{R}^3$ . Nous notons  $u = \frac{\partial \varphi}{\partial x}(a)$ ,  $v = \frac{\partial \varphi}{\partial y}(a)$  et  $w = \frac{\partial \varphi}{\partial z}(a)$ .

En utilisant l'expression du lemme 10.41 à propos du produit mixte, la formule 20.34 donne

$$\det(d\varphi_a^* \circ d\varphi_a) = \det \begin{pmatrix} \|u\|^2 & u \cdot v & u \cdot w \\ v \cdot u & \|v\|^2 & v \cdot w \\ w \cdot u & w \cdot v & \|w\|^2 \end{pmatrix} = |(u \times v) \cdot w|. \quad (20.37)$$

Donc

$$\det(d\varphi_a^* \circ d\varphi_a) = \left| \left( \frac{\partial \varphi}{\partial x}(a) \times \frac{\partial \varphi}{\partial y}(a) \right) \cdot \frac{\partial \varphi}{\partial z}(a) \right|. \quad (20.38)$$

## 20.3 Intégrale sur une variété

### 20.3.1 Mesure sur une carte

Nous considérons dans cette section uniquement des variétés  $M$  de dimension 2 dans  $\mathbb{R}^3$ . Une particularité de  $\mathbb{R}^3$  (par rapport aux autres  $\mathbb{R}^n$ ) est qu'il existe le produit vectoriel.

Si  $v, w \in \mathbb{R}^3$ , alors le vecteur  $v \times w$  est un vecteur normal au plan décrit par  $v$  et  $w$  qui jouit de l'importante propriété suivante :

$$\text{aire du parallélogramme} = \|v \times w\|. \quad (20.39)$$

L'aire du parallélogramme construit sur  $v$  et  $w$  est donnée par la norme du produit vectoriel. Afin de donner une mesure infinitésimale en un point  $p \in M$ , nous voudrions prendre deux vecteurs tangents à  $M$  en  $p$ , et puis considérer la norme de leur produit vectoriel. Cette idée se heurte à la question du choix des vecteurs tangents à considérer.

Dans  $\mathbb{R}^2$ , le choix est évident : nous choisissons  $e_x$  et  $e_y$ , et nous avons  $\|e_x \times e_y\| = 1$ . L'idée est donc de choisir une carte  $F: W \rightarrow F(W)$  autour du point  $p = F(w)$ , et de choisir les vecteurs tangents qui correspondent à  $e_x$  et  $e_y$  via la carte, c'est-à-dire les vecteurs

$$\frac{\partial F}{\partial x}(w), \quad \text{et} \quad \frac{\partial F}{\partial y}(w). \quad (20.40)$$

L'élément infinitésimal de surface sur  $M$  au point  $p = F(w)$  est alors défini par

$$d\sigma_F = \left\| \frac{\partial F}{\partial x}(w) \times \frac{\partial F}{\partial y}(w) \right\| dw, \quad (20.41)$$

et si la partie  $A \subset M$  est entièrement contenue dans  $F(W)$ , nous définissons la **mesure** de  $A$  par

$$\mu_2(A) = \int_{F^{-1}(A)} d\sigma_F = \int_{F^{-1}(A)} \left\| \frac{\partial F}{\partial x}(w) \times \frac{\partial F}{\partial y}(w) \right\| dw. \quad (20.42)$$

#### Remarque 20.12.

Afin que cette définition ait un sens, nous devons prouver qu'elle ne dépend pas du choix de la carte  $F$ . En effet, les vecteurs  $\frac{\partial F}{\partial x}$  et  $\frac{\partial F}{\partial y}$  dépendent de la carte  $F$ , donc leur produit vectoriel (et sa norme) dépendent également de la carte  $F$  choisie. Il faudrait donc un petit miracle pour que le nombre  $\mu_2(A)$  donné par (20.42) soit indépendant du choix de  $F$ . Nous allons bientôt voir comme cas particulier du théorème 20.16 que c'est en fait le cas. C'est-à-dire que si  $F$  et  $\tilde{F}$  sont deux cartes qui contiennent  $A$ , alors

$$\int_{F^{-1}(A)} d\sigma_F = \int_{\tilde{F}^{-1}(A)} d\sigma_{\tilde{F}}. \quad (20.43)$$

### 20.3.1.1 Exemple : la mesure de la sphère

**Proposition 20.13** ([1]).

L'aire de la sphère  $S^2 \equiv x^2 + y^2 + z^2 = R^2$  est  $4\pi R^2$ .

*Démonstration.* Nous allons calculer l'aire de la demi-sphère. L'application  $F: B((0,0), R) \rightarrow \mathbb{R}^3$  donnée par

$$F(x, y) = \begin{pmatrix} x \\ y \\ \sqrt{R^2 - x^2 - y^2} \end{pmatrix} \quad (20.44)$$

est une carte pour une demi-sphère. Ses dérivées partielles sont

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ -\frac{x}{\sqrt{R^2 - x^2 - y^2}} \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ -\frac{y}{\sqrt{R^2 - x^2 - y^2}} \end{pmatrix}. \quad (20.45)$$

Le produit vectoriel de ces deux vecteurs tangents donne

$$\frac{\partial F}{\partial x}(x, y) \times \frac{\partial F}{\partial y}(x, y) = \frac{x}{\alpha} e_1 + \frac{y}{\alpha} e_2 + e_3 \quad (20.46)$$

où  $\alpha = \sqrt{R^2 - x^2 - y^2}$ . En calculant la norme, nous trouvons

$$\left\| \frac{\partial F}{\partial x}(x, y) \times \frac{\partial F}{\partial y}(x, y) \right\| = \sqrt{\frac{R^2}{R^2 - x^2 - y^2}}, \quad (20.47)$$

et en passant aux coordonnées polaires, nous écrivons l'intégrale (20.42) sous la forme

$$\int_B \|\partial_x F \times \partial_y F\| = \int_0^{2\pi} d\theta \int_0^R r \sqrt{\frac{R^2}{R^2 - x^2 - y^2}} dr = 2\pi R^2, \quad (20.48)$$

qui est bien la mesure de la demi-sphère.  $\square$

### 20.3.2 Intégrale sur une carte

Nous pouvons maintenant définir l'intégrale d'une fonction sur une carte de la variété  $M$ .

**Définition 20.14.**

Soit  $F: W \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , une carte pour une variété  $M$ . Soit  $A$ , une partie de  $F(W)$  telle que  $A = F(B)$  où  $B \subset W$  est mesurable. Soit encore  $f: A \rightarrow \mathbb{R}$ , une fonction continue. L'intégrale de  $f$  sur  $A$  est le nombre

$$\int_A f = \int_A f d\sigma_F = \int_{F^{-1}(A)} (f \circ F)(w) \left\| \frac{\partial F}{\partial x}(w) \times \frac{\partial F}{\partial y}(w) \right\| dw \quad (20.49)$$

**Remarque 20.15.**

L'intégrale (20.49) n'est pas toujours bien définie. Étant donné que  $F$  est  $C^1$  et que  $f$  est continue, l'intégrande est continue. L'intégrale sera donc bien définie par exemple lorsque  $B$  est borné et si la fermeture  $\bar{A}$  est un compact contenu dans  $F(W)$ .

Le théorème suivant montre que le travail que nous avons fait jusqu'à présent ne dépend en fait pas du choix de carte  $F$  effectué.

**Théorème 20.16.**

Soient  $F: W \rightarrow F(W)$  et  $\tilde{F}: \tilde{W} \rightarrow \tilde{F}(\tilde{W})$ , deux cartes de la variété  $M$ . Soit une partie  $A \subset F(W) \cap \tilde{F}(\tilde{W})$  telle que  $A = F(B)$  avec  $B \subset W$  mesurable. Alors  $A = \tilde{F}(\tilde{B})$  avec  $\tilde{B} \subset \tilde{W}$  mesurable.

Si  $f$  est une fonction continue, et si  $\int_A f d\sigma_F$  existe, alors  $\int_A f d\sigma_{\tilde{F}}$  existe et

$$\int_A f d\sigma_F = \int_A f d\sigma_{\tilde{F}}. \quad (20.50)$$

### 20.3.3 Exemples

Intégrons la fonction  $f(x, y, z)$  sur le carré  $K = ]0, 1[ \times ]0, 2[ \times \{1\}$ . La première carte que nous pouvons utiliser est

$$\begin{aligned} F: ]0, 1[ \times ]0, 2[ &\rightarrow K \\ (x, y) &\mapsto (x, y, 1). \end{aligned} \quad (20.51)$$

Nous trouvons aisément les vecteurs tangents qui forment l'élément de surface :

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad (20.52)$$

donc  $d\sigma_F = 1 \cdot dx dy$ , et

$$\int_K f d\sigma_F = \int_{]0,1[ \times ]0,2[} f(x, y, 1) \cdot 1 \cdot dx dy. \quad (20.53)$$

Nous pouvons également utiliser la carte

$$\begin{aligned} \tilde{F}: ]0, \frac{1}{2}[ \times ]0, 6[ &\rightarrow K \\ (\tilde{x}, \tilde{y}) &\mapsto (2\tilde{x}, \frac{\tilde{y}}{3}, 1). \end{aligned} \quad (20.54)$$

Les vecteurs tangents sont maintenant

$$\frac{\partial \tilde{F}}{\partial \tilde{x}} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial \tilde{F}}{\partial \tilde{y}} = \begin{pmatrix} 0 \\ 1/3 \\ 0 \end{pmatrix}, \quad (20.55)$$

et nous avons donc  $d\sigma_{\tilde{F}} = \|\frac{2}{3}e_3\| = \frac{2}{3}$ . Cette fois, l'intégrale de  $f$  sur  $K$  s'écrit

$$\int_K f d\sigma_{\tilde{F}} = \int_{]0, \frac{1}{2}[ \times ]0, 6[} f(2\tilde{x}, \frac{\tilde{y}}{3}, 1) \cdot \frac{2}{3} \cdot d\tilde{x} d\tilde{y}. \quad (20.56)$$

Conformément au théorème 20.16, cette dernière intégrale est égale à l'intégrale (20.53) parce qu'il s'agit juste d'un changement de variable.

### 20.3.4 Orientation

Soient  $F: W \rightarrow F(W)$  et  $\tilde{F}: \tilde{W} \rightarrow \tilde{F}(\tilde{W})$ , deux cartes de la variété  $M$ . Nous pouvons considérer la fonction  $h = \tilde{F}^{-1} \circ F$ , définie uniquement sur l'intersection des cartes :

$$h: F^{-1}(F(W) \cap \tilde{F}(\tilde{W})) \rightarrow \tilde{F}^{-1}(F(W) \cap \tilde{F}(\tilde{W})). \quad (20.57)$$

Nous disons que  $F$  et  $\tilde{F}$  ont même **orientation** si

$$J_h(w) > 0 \quad (20.58)$$

pour tout  $w \in F^{-1}(F(W) \cap \tilde{F}(\tilde{W}))$ .

Considérons les deux cartes suivantes pour le même carré :

$$\begin{aligned} F: ]0, 1[ \times ]0, 1[ &\rightarrow \mathbb{R}^3 \\ (x, y) &\mapsto (x, y, 0) \end{aligned} \quad (20.59)$$

et

$$\begin{aligned} \tilde{F}: ]0, \frac{1}{2}[ \times ]0, \frac{1}{3}[ &\rightarrow \mathbb{R}^3 \\ (x, y) &\mapsto (2x, 3y, 0) \end{aligned} \quad (20.60)$$

Ici,  $h(x, y) = \left(\frac{x}{2}, \frac{y}{3}\right)$  et nous avons  $J_h = \frac{1}{6} > 0$ . Ces deux cartes ont même orientation. Notez que

$$\frac{\partial F}{\partial x} \times \frac{\partial F}{\partial y} = e_3, \quad (20.61)$$

tandis que

$$\frac{\partial \tilde{F}}{\partial x} \times \frac{\partial \tilde{F}}{\partial y} = 6e_3. \quad (20.62)$$

Les vecteurs normaux à le paramétrage pointent dans le même sens.

Si par contre nous prenons le paramétrage

$$\begin{aligned} G: ]0, 1[ \times ]0, 1[ &\rightarrow \mathbb{R}^3 \\ (x, y) &\mapsto (x, (1 - y), 0), \end{aligned} \quad (20.63)$$

nous avons

$$\frac{\partial G}{\partial x} \times \frac{\partial G}{\partial y} = -e_3, \quad (20.64)$$

et si  $g = G^{-1} \circ F$ , alors  $J_g = -1$ .

L'orientation d'une carte montre donc si le vecteur normal à la surface pointe d'un côté ou de l'autre de la surface.

DEFooSWREooNdQpdA

**Définition 20.17** (Variété orientable).

Une variété  $M$  est **orientable** si il existe un atlas de  $M$  tel que deux cartes quelconques ont toujours même orientation. Une variété est **orientée** lorsque qu'un tel choix d'atlas est fait.

**Proposition 20.18.**

Soit  $M$ , une variété orientable et un atlas orienté  $\{F_i: W_i \rightarrow \mathbb{R}^3\}$ . Alors le vecteur unitaire

$$\frac{\frac{\partial F}{\partial x}(x, y) \times \frac{\partial F}{\partial y}(x, y)}{\left\| \frac{\partial F}{\partial x}(x, y) \times \frac{\partial F}{\partial y}(x, y) \right\|} \quad (20.65)$$

ne dépend pas du choix de  $F$  parmi les  $F_i$ .

*Démonstration.* Considérons deux cartes  $F_1$  et  $F_2$ , ainsi que l'application  $h = F_2^{-1} \circ F_1$ . Écrivons le vecteur  $\partial_x F_1 \times \partial_y F_1$  en utilisant  $F_1 = F_2 \circ h$ . D'abord, par la règle de dérivation de fonctions composées,

$$\frac{\partial(F_2 \circ h)}{\partial x} = \frac{\partial F_2}{\partial x} \frac{\partial h_1}{\partial x} + \frac{\partial F_2}{\partial y} \frac{\partial h_2}{\partial x}. \quad (20.66)$$

Après avoir fait le même calcul pour  $\frac{\partial(F_2 \circ h)}{\partial y}$ , nous pouvons écrire

$$\partial_x(F_2 \circ h) \times \partial_y(F_2 \circ h) = (\partial_x h_1 \partial_x F_2 + \partial_x h_2 \partial_y F_2) \times (\partial_y h_1 \partial_x F_2 + \partial_y h_2 \partial_y F_2). \quad (20.67)$$

Dans cette expression, les facteurs  $\partial_i h_j$  sont des nombres, donc ils se factorisent dans les produits vectoriels. En tenant compte du fait que  $\partial_x F_2 \times \partial_x F_2 = 0$  et  $\partial_y F_2 \times \partial_y F_2 = 0$ , ainsi que de l'antisymétrie du produit vectoriel, l'expression se réduit à

$$\left( \frac{\partial F_2}{\partial x} \times \frac{\partial F_2}{\partial y} \right) (\partial_x h_1 \partial_y h_2 - \partial_x h_2 \partial_y h_1). \quad (20.68)$$

Par conséquent,

$$\frac{\partial F_1}{\partial x} \times \frac{\partial F_1}{\partial y} = \frac{\partial(F_2 \circ h)}{\partial x} \times \frac{\partial(F_2 \circ h)}{\partial y} = \left( \frac{\partial F_2}{\partial x} \times \frac{\partial F_2}{\partial y} \right) \det J_h. \quad (20.69)$$

Donc, tant que  $J_h$  est positif, les vecteurs unitaires correspondants au membre de gauche et de droite sont égaux.  $\square$

**Corolaire 20.19.**

Si nous avons choisi un atlas orienté pour la variété  $M$ , nous avons une fonction continue  $G: M \rightarrow \mathbb{R}^3$  telle que  $\|G(p)\| = 1$  pour tout  $p \in M$ . Cette fonction est donnée par

$$G(F(x, y)) = \frac{\frac{\partial F}{\partial x}(x, y) \times \frac{\partial F}{\partial y}(x, y)}{\left\| \frac{\partial F}{\partial x}(x, y) \times \frac{\partial F}{\partial y}(x, y) \right\|} \quad (20.70)$$

sur l'image de la carte  $F$ .

*Démonstration.* La fonction  $G$  est construite indépendamment sur chaque carte  $F(W)$  en utilisant la formule (20.70). Cette fonction est une fonction bien définie sur tout  $M$  parce que nous venons de démontrer que sur  $F_1(W_1) \cap F_2(W_2)$ , les fonctions construites à partir de  $F_1$  et à partir de  $F_2$  sont égales.  $\square$

Il est possible que prouver, bien que cela soit plus compliqué, que la réciproque est également vraie.

**Proposition 20.20.**

Une variété  $M$  de dimension 2 dans  $\mathbb{R}^3$  est orientable si et seulement si il existe une fonction continue  $G: M \rightarrow \mathbb{R}^3$  telle que pour tout  $p \in M$ , le vecteur  $G(p)$  soit de norme 1 et normal à  $M$  au point  $p$ .

**20.4 Longueur, aire, volumes etc.**

Grâce à la mesure de Lebesgue (définition 14.249), nous avons la définition d'aires dans  $\mathbb{R}^2$  et de volumes dans  $\mathbb{R}^3$ . Dans tout ce qui suit, nous considérons toujours la tribu de Lebesgue, la mesure de Lebesgue, et l'intégrale de Lebesgue correspondante.

**Définition 20.21.**

L'*aire* de la partie mesurable  $S$  de  $\mathbb{R}^2$  est le nombre

$$\int_{\mathbb{R}^2} \mathbb{1}_S \quad (20.71)$$

au sens de l'intégrale pour la mesure de Lebesgue de la fonction caractéristique de la partie  $S$ .

**Définition 20.22.**

Le *volume* de la partie mesurable  $V$  de  $\mathbb{R}^3$  est le nombre

$$\int_{\mathbb{R}^3} \mathbb{1}_V \quad (20.72)$$

au sens de l'intégrale pour la mesure de Lebesgue de la fonction caractéristique de la partie  $V$ .

Ceci est bien beau, mais ne permet pas de définir l'aire d'une surface dans  $\mathbb{R}^3$ , ni une longueur dans  $\mathbb{R}^2$ . Nous n'avons pas encore défini ce que nous appelons une surface dans  $\mathbb{R}^3$ , mais selon toute définition raisonnable, si  $S$  en est une, elle sera négligeable au sens de la mesure de Lebesgue dans  $\mathbb{R}^3$  et nous aurons toujours

$$\int_{\mathbb{R}^3} \mathbb{1}_S = 0. \quad (20.73)$$

L'objet de ce chapitre sera de donner un sens aux notions de longueurs dans  $\mathbb{R}^2$ , de surfaces dans  $\mathbb{R}^3$  et d'y définir des intégrales permettant de définir longueurs et aires.

**20.4.1 Quelques aires faciles**

Nous nous souvenons de la proposition 10.45 qui donnait une bonne propriété du produit vectoriel dans  $\mathbb{R}^3$ . Nous en donnons une autre.

LEMooVHGKooDjefOL

**Lemme 20.23.**

*L'aire d'une droite dans  $\mathbb{R}^2$  est nulle.*

*Démonstration.* Nous considérons la mesure le Lebesgue  $\lambda$  sur  $\mathbb{R}$  et une droite  $A$  dans  $\mathbb{R}^2$ .

La mesure dans  $\mathbb{R}^2$  est donné par la définition 14.249 qui demande de calculer les intégrales (14.754) :

$$(\lambda \otimes \lambda)(A) = \int_{\mathbb{R}} \lambda(A_2(x)) d\lambda(x) = \int_{\mathbb{R}} \lambda(A_1(y)) d\lambda(y) \quad (20.74)$$

EQooUTAHooUyESco

où  $A_2(x) = \{y \in \mathbb{R} \text{ tel que } (x, y) \in A\}$  et  $A_1(y) = \{x \in \mathbb{R} \text{ tel que } (x, y) \in A\}$ .

En vertu du lemme 12.152, cette droite est soit d'équation  $y = ax + b$ , soit d'équation  $x = a$ .

- (1) **Droite**  $y = ax + b$  Dans ce cas,  $A_2(x) = \{ax + b\}$ . C'est un singleton et nous avons  $\lambda(A_2(x)) = 0$  pour tout  $x$ . Nous avons donc  $(\lambda \otimes \lambda)(A) = 0$ .
- (2) **Droite**  $x = a$  Nous utilisons la seconde possibilité laissée par l'égalité (20.74). L'ensemble  $A_1$  est facile à déterminer :  $A_1(y) = \{a\}$  pour tout  $y$ . Donc  $\lambda(A_1(y)) = 0$  pour tout  $y$ , et l'intégrale est nulle.

□

**Remarque 20.24.**

Dans la preuve de 20.23, nous aurions pu faire le cas  $x = a$  en utilisant la première formule. Dans ce cas nous serions partis de

$$A_2(x) = \begin{cases} \mathbb{R} & \text{si } x = a \\ \emptyset & \text{sinon,} \end{cases} \quad (20.75)$$

et donc de

$$\lambda(A_2(x)) = \begin{cases} \infty & \text{si } x = a \\ 0 & \text{sinon.} \end{cases} \quad (20.76)$$

Le lemme 14.172 nous dit que l'intégrale de  $A_2$  sur  $\mathbb{R}$  est nulle.

DEFooDTFCooCTdaDO

**Définition 20.25.**

Soient  $a, u_1$  et  $u_2$  dans  $\mathbb{R}^2$ . Le **parallélogramme** basé en  $a$  formé sur  $u_1$  et  $u_2$  est l'ensemble

$$\{a + su_1 + tu_2 \text{ tel que } 0 \leq t \leq 1, 0 \leq s \leq 1\}. \quad (20.77)$$

Assez souvent, nous supposeront que  $u_1$  et  $u_2$  ne sont pas colinéaires.

PROPOoAVVNooOOISzr

**Proposition 20.26.**

L'aire<sup>5</sup> du parallélogramme<sup>6</sup> basé en  $a$  et formé sur  $u$  et  $v$  (que nous supposons n'être pas colinéaires) est donnée par

$$\|u \times v\| = |u_1 v_2 - v_1 u_2|. \quad (20.78)$$

*Démonstration.* En vertu de la définition 20.21 d'une aire, le nombre à calculer est  $\int_D 1 d\lambda$  où

$$D_0 = \{a + xu + yv \text{ tel que } x \in [0, 1], y \in [0, 1]\}. \quad (20.79)$$

En posant

$$D = \{a + xu + yv \text{ tel que } x \in ]0, 1[, y \in ]0, 1[ \}, \quad (20.80)$$

nous avons

$$\int_{D_0} d\lambda = \int_D d\lambda, \quad (20.81)$$

parce que ce que nous avons enlevé sont des segments de droites alors que les droites sont de mesure nulle (lemme 20.23).

5. Définition 20.21.

6. Définition 20.25.

Nous considérons la paramétrisation

$$\begin{aligned}\varphi: ]0, 1[ \times ]0, 1[ &\rightarrow D \\ (x, y) &\mapsto a + xu + yv.\end{aligned}\tag{20.82}$$

Pour le théorème 14.282, nous devons montrer que  $\phi$  est un  $C^1$ -difféomorphisme. Comme nous sommes un peu fatigués, nous allons seulement prouver  $\phi$  est injective<sup>7</sup>. Supposons que  $\phi(x_1, x_2) = \phi(y_1, y_2)$ . Alors

$$(x_1 - x_2)u + (y_1 - y_2)v = 0.\tag{20.83}$$

Si  $y_1 - y_2 = 0$ , alors  $x_1 - x_2 = 0$  et on est bon. Sinon,

$$v = -\frac{x_1 - x_2}{y_1 - y_2}u\tag{20.84}$$

Vu que les vecteurs  $u$  et  $v$  ne sont pas colinéaires, nous en déduisons que  $x_1 - x_2 = 0$  et on est encore bon.

Nous utilisons la formule (14.860) de changement de variable avec la fonction  $\phi$ . Le jacobien à calculer est

$$J = \det \begin{pmatrix} (\partial_x \phi_1)(x, y) & (\partial_x \phi_2)(x, y) \\ (\partial_y \phi_1)(x, y) & (\partial_y \phi_2)(x, y) \end{pmatrix} = \det \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix} = |u_1 v_2 - v_1 u_2|.\tag{20.85}$$

Le lien avec le produit vectoriel est un petit abus de notation : il s'agit du produit vectoriel entre  $(u_1, u_2, 0)$  et  $(v_1, v_2, 0)$  qui peut être obtenu en utilisant directement la formule de définition (10.104).  $\square$

## 20.5 Autres théorèmes de points fixes

En termes de théorème de points fixes nous avons déjà vu le théorème de Picard 17.34. Voir aussi le thème 65.

### 20.5.1 Brouwer

#### Proposition 20.27.

Soit  $f: [a, b] \rightarrow [a, b]$  une fonction continue. Alors  $f$  accepte un point fixe.

*Démonstration.* En effet si nous considérons  $g(x) = f(x) - x$  alors nous avons  $g(a) = f(a) - a \geq 0$  et  $g(b) = f(b) - b \leq 0$ . Si  $g(a)$  ou  $g(b)$  est nul, la proposition est démontrée ; nous supposons donc que  $g(a) > 0$  et  $g(b) < 0$ . La proposition découle à présent du théorème des valeurs intermédiaires 9.95.  $\square$

#### Exemple 20.28.

La fonction  $x \mapsto \cos(x)$  est continue entre  $[-1, 1]$  et  $[-1, 1]$ . Elle admet donc un point fixe. Par conséquent il existe (au moins) une solution à l'équation  $\cos(x) = x$ .  $\triangle$

#### Proposition 20.29 (Théorème de rétraction[645]).

Il n'existe pas d'applications  $g: \overline{B(0, 1)} \rightarrow \partial B(0, 1)$  qui soit de classe  $C^1$  et qui soit l'identité sur  $\partial B(0, 1)$ .

*Démonstration.* En plein d'étapes.

(1) Définition de l'application  $F$  Nous définissons la fonction

$$\begin{aligned}F: \overline{B(0, 1)} \times [0, 1] &\rightarrow \overline{B(0, 1)} \\ (x, t) &\mapsto (1 - t)x + tg(x).\end{aligned}\tag{20.86}$$

7. Pour le reste, écrivez l'inverse explicitement, et prouvez que c'est  $C^1$ . Si ça pose problème, écrivez-moi parce que je n'ai pas vérifié.

Lorsque  $t$  est fixé, nous notons

$$\begin{aligned} F_t: \overline{B(0,1)} &\rightarrow \overline{B(0,1)} \\ x &\mapsto F(x,t). \end{aligned} \quad (20.87)$$

Nous avons  $F_0 = \text{Id}|_{\overline{B(0,1)}}$  et  $F_1 = g$ . De plus, nous avons  $F(x,t) = x$  pour tout  $x \in \partial B(0,1)$  et pour tout  $t \in [0,1]$ .

(2) **définition de  $V$**  Nous posons<sup>8</sup>

$$\begin{aligned} V: [0,1] &\rightarrow \mathbb{R} \\ t &\mapsto \int_{B(0,1)} \det(dF_t)_x dx. \end{aligned} \quad (20.88)$$

En ce qui concerne la différentielle de  $F_t$ , nous avons  $F_t(x) = (1-t)x + tg(x)$ . La différentielle de  $x \mapsto x$  est  $v \mapsto v$  (lemma 10.294), et donc

$$(dF_t)_x = (1-t)\text{Id} + tdg_x. \quad (20.89)$$

Le déterminant de cela est un polynôme en  $t$ .

(3)  $V(1) = 0$  Vu que  $g(x) \in \partial B(0,1)$ , nous avons  $g(x) \cdot g(x) = \|g(x)\|^2 = 1$ . En utilisant la proposition 10.310, pour tout  $h \in \mathbb{R}^n$  nous avons :

$$2g(x) \cdot (dg_x)(h) = 0. \quad (20.90)$$

Autrement dit, l'image de  $dg_x$  est dans  $g(x)^\perp$ , qui est de dimension plus petite que  $n$  (parce que  $g(x) \neq 0$ ). Donc  $dg_x$  n'est pas de rang maximum ; bref : son déterminant est nul.

Vu que  $F_1(x) = F(1,x) = g(x)$ , nous avons bien  $V(1) = 0$ .

(4)  **$\det((dF_t)_x) > 0$  si  $t$  est assez petit**

Supposons le contraire : pour tout  $n \in \mathbb{N}$ , il existe  $t_n \in ]0, 1/n]$  et  $x_n \in \overline{B(0,1)}$  tels que  $\det((dF_{t_n})_{x_n}) \leq 0$ . La suite  $(x_n)$  étant dans le compact  $\overline{B(0,1)}$ , elle admet une sous-suite convergente : il existe une application strictement croissante  $\phi: \mathbb{N} \rightarrow \mathbb{N}$  et un  $x \in \overline{B(0,1)}$  tels que  $x_{\phi(n)} \rightarrow x$ . Notez au passage que  $t_{\phi(n)} \rightarrow 0$ .

Pour tout  $n$  nous avons

$$\det((dF_{t_{\phi(n)}})_{x_{\phi(n)}}) \leq 0. \quad (20.91)$$

Vu que l'application  $x \mapsto \det((dF_t)_x)$  est continue en  $t$  et en  $x$ , nous pouvons passer à la limite dans (20.91) :

$$\det((dF_0)_x) \leq 0. \quad (20.92)$$

Mais comme  $dF_0 = \text{Id}$ , nous avons en réalité  $\det((dF_0)_x) > 0$ . Contradiction. Nous en déduisons que il existe  $m_1 > 0$  tel que pour tout  $t < m_1$  et pour tout  $x \in \overline{B(0,1)}$ , nous avons

$$\det((dF_t)_x) > 0. \quad (20.93)$$

(5)  **$F_t$  est injective si  $t$  est assez petit**

Soient  $t < 1/N$  et  $x, y \in \overline{B(0,1)}$  tels que  $F(x,t) = F(y,t)$ . Nous avons

$$0 = F(x,t) - F(y,t) = (1-t)(x-y) + t(g(x) - g(y)). \quad (20.94)$$

En passant aux normes et en utilisant le théorème des accroissements finis 10.324,

$$(1-t)\|x-y\| = t\|g(x) - g(y)\| \leq tM\|x-y\| \quad (20.95)$$

8. L'intégrale est définie en (14.539), et la mesure de Lebesgue en 14.249. Mais en principe, ça fait longtemps que vous êtes à l'aise avec ça.



où  $M = \sup_{x \in \overline{B(0,1)}} \|dg_x\|$ . En simplifiant par  $\|x - y\|$ , nous trouvons  $tM/(1-t) \geq 1$  ou encore

$$t \geq \frac{1}{M+1}. \quad (20.96)$$

Autrement dit, dès que  $t < 1/(M+1)$ , l'application  $F_t$  est injective.

(6)  **$t$  est assez petit**

Nous posons

$$\alpha = \min\{m_1, \frac{1}{M+1}\}. \quad (20.97)$$

Dès que  $t < \alpha$ ,

- $F_t$  est injective,
- $\det((dF_t)_x) > 0$ .

Cela étant dit, nous continuons la preuve.

(7) **Un peu de relations ensemblistes**

Nous considérons  $t < \alpha$  et  $x \in \partial B(0, 1)$ . Nous avons alors  $F(x, t) = x$ . Nous avons alors

$$F(\partial B(0, 1), t) \cup B(0, 1) = \overline{B(0, 1)}, \quad (20.98)$$

qui est fermé dans  $\mathbb{R}^n$ .

Vu que  $F_t$  est injective, nous pouvons utiliser la formule ensembliste (1.14) : SUBEQSooLWJLooBEVkro

$$F(B(0, 1), t) = F(\overline{B(0, 1)}, t) \setminus F(\partial B(0, 1), t) \quad (20.99a)$$

$$= F(\overline{B(0, 1)}, t) \setminus \partial B(0, 1) \quad (20.99b)$$

$$= F(\overline{B(0, 1)}, t) \cap B(0, 1), \quad (20.99c)$$

parce que  $X \setminus \partial B(0, 1) = X \cap B(0, 1)$ .

(8)  **$F(B(0, 1), t)$  est ouvert**

L'application  $F$  est de classe  $C^1$  et injective. Donc  $(dF_t)_x$  est continue et injective (proposition 10.279). Soit  $x \in B(0, 1)$ . Par le théorème d'inversion locale 17.47 il existe un ouvert  $V_x \subset B(0, 1)$  et un ouvert  $W_x$  de  $\mathbb{R}^n$  tels que  $F_t: V_x \rightarrow W_x$  soit une bijection. Nous avons donc

$$F_t(B(0, 1)) = \bigcup_{x \in B(0, 1)} F_t(V_x) = \bigcup_{x \in B(0, 1)} W_x. \quad (20.100) \quad \text{EQooDCMAooVJnhxy}$$

Une union d'ouverts étant ouverte, nous en déduisons que  $F(B(0, 1), t)$  est ouvert.

(9) **Un peu de topologie induite** Nous considérons l'espace  $B(0, 1)$  avec la topologie induite de  $\mathbb{R}^n$ .

(9a) **C'est fermé**

La partie  $F(\overline{B(0, 1)}, t)$  est l'image d'un compact par une fonction continue. C'est donc un compact (et donc fermé) par le théorème 7.274. Du coup la proposition 7.30 dit que la partie

$$F(B(0, 1), t) = F(\overline{B(0, 1)}, t) \cap B(0, 1) \quad (20.101)$$

est fermée dans  $B(0, 1)$ .

(9b) **C'est ouvert**

La partie  $F(B(0, 1), t)$  est ouverte (voir (20.100)) et contenue dans  $B(0, 1)$  (voir (20.99)). Elle est donc un ouvert pour la topologie de  $B(0, 1)$ .

(9c) **Et donc c'est ...**

La partie  $F(B(0, 1), t)$  est ouvert et fermée dans  $B(0, 1)$ . Par connexité, la proposition 7.105(4) conclut que

$$F(B(0, 1), t) = B(0, 1). \quad (20.102)$$

(10) **Volume de la boule**

Nous sommes toujours avec  $t < \alpha$ . Nous avons donc  $\det((dF_t)_x) > 0$  pour tout  $x$ . Nous calculons le volume de la boule  $B(0, 1)$  en introduisant la fonction constante 1. En utilisant le changement de variables<sup>9</sup>  $F_t: B(0, 1) \rightarrow B(0, 1)$  :

$$\int_{B(0,1)} 1dx = \int_{F_t^{-1}(B(0,1))} \underbrace{(1 \circ F_t)(x)}_{=1} |\det((dF_t)_x)| dx \quad (20.103a)$$

$$= \int_{B(0,1)} \det((dF_t)_x) dx \quad (20.103b)$$

$$= V(t). \quad (20.103c)$$

Nous venons de prouver que sur  $t \in [0, \alpha[$ , la fonction  $V(t)$  est en réalité une constante égale au volume de la boule unité.

(11) **Et enfin...**

Vu que  $V$  est un polynôme et que  $V$  est constant sur un intervalle,  $V$  est constante partout. En particulier,  $V(1)$  est encore égal au volume de la boule. Ah ben non parce que nous avons démontré plus haut que  $V(1) = 0$ .

Contradiction, fin de l'histoire.

□

PropDRpYwv

**Proposition 20.30** (Brouwer dans  $\mathbb{R}^n$  version  $C^\infty$ ).

Toute application  $f: \overline{B(0, 1)} \rightarrow \overline{B(0, 1)}$  de classe  $C^\infty$  admet un point fixe.

*Démonstration.* Supposons que  $f$  ne possède pas de points fixes. Alors pour tout  $x \in \overline{B(0, 1)}$  nous considérons la demi-droite issue de  $f(x)$  et passant par  $x$  (cette demi-droite existe parce que  $x$  et  $f(x)$  sont supposés distincts). Cette demi-droite intersecte  $\partial B(0, 1)$  en un point que nous appelons  $g(x)$ . Prouvons que cette fonction est  $C^k$  dès que  $f$  est  $C^k$  ( $y$  compris avec  $k = \infty$ ).

L'application

$$g: \overline{B(0, 1)} \rightarrow \partial B(0, 1) \quad (20.104)$$

est construite comme solution du système

$$\begin{cases} g(x) - f(x) = \lambda(x - f(x)) & (20.105a) \\ \|g(x)\|^2 = 1 & (20.105b) \\ \lambda \geq 0. & (20.105c) \end{cases}$$

En substituant nous obtenons l'équation

$$P_x(\lambda) = \|\lambda(x - f(x)) + f(x)\|^2 - 1 = 0, \quad (20.106)$$

ou encore

$$\lambda^2 \|x - f(x)\|^2 + 2\lambda(x - f(x)) \cdot f(x) + \|f(x)\|^2 - 1 = 0. \quad (20.107)$$

En tenant compte du fait que  $\|f(x)\| < 1$  (parce que les images de  $f$  sont dans  $\mathcal{B}$ ), nous trouvons que  $P_x(0) \leq 0$  et  $P_x(1) \leq 0$ . De même  $\lim_{\lambda \rightarrow \infty} P_x(\lambda) = +\infty$ . Par conséquent le polynôme de second degré  $P_x$  a exactement deux racines distinctes  $\lambda_1 \leq 0$  et  $\lambda_2 \geq 1$ . La racine que nous cherchons est la seconde. Le discriminant est strictement positif, donc pas besoin d'avoir peur de la racine dans

$$\lambda(x) = \frac{-(x - f(x)) \cdot f(x) + \sqrt{\Delta_x}}{\|x - f(x)\|^2} \quad (20.108)$$

où

$$\Delta_x = 4\left((x - f(x)) \cdot f(x)\right)^2 - 4\|x - f(x)\|^2(\|f(x)\|^2 - 1). \quad (20.109)$$

Notons que la fonction  $\lambda(x)$  est  $C^k$  dès que  $f$  est  $C^k$ ; et en particulier elle est  $C^\infty$  si  $f$  l'est.

En résumé la fonction  $g$  ainsi définie vérifie deux propriétés :

---

9. Théorème 14.282.

- (1)  $g: \overline{B(0,1)} \rightarrow \overline{B(0,1)}$
- (2) elle est  $C^\infty$ ;
- (3) elle est l'identité sur  $\partial B(0,1)$ .

Cela viole le théorème de contraction 20.29. Contradiction, fin de la preuve.  $\square$

### 20.31.

Voici comment on démontre le théorème de rétraction dans le cas  $C^\infty$  en utilisant Stokes.

La suite de la preuve consiste à montrer qu'une telle application  $g: B \rightarrow \partial B$  ne peut pas exister<sup>10</sup>. Nous considérons une forme de volume  $\omega$  sur  $\partial B$  : l'intégrale de  $\omega$  sur  $\partial B$  est l'aire de  $\partial B$  qui est non nulle. Nous avons alors la contradiction suivante :

$$0 < \int_{\partial B} \omega = \int_{\partial B} g^* \omega = \int_B d(g^* \omega) = \int_B g^*(d\omega) = 0 \quad (20.110)$$

Justifications :

- L'intégrale  $\int_{\partial B} \omega$  est l'aire de  $\partial B$  et est donc strictement positive.
- La fonction  $g$  est l'identité sur  $\partial B$ . Nous avons donc  $\omega = g^* \omega$ .
- Le lemme 27.95.
- La forme  $\omega$  est de volume, par conséquent de degré maximum et  $d\omega = 0$ .

Un des points délicats pour démontrer le théorème de Brouwer en version continue est de se ramener au cas de fonctions  $C^\infty$ . Pour la régularisation par convolution, voir [646] ; pour celle utilisant le théorème de Weierstrass, voir [647].

ThoRGjGdO

**Théorème 20.32** (Brouwer dans  $\mathbb{R}^n$ , version continue).

Toute application continue<sup>11</sup>  $\overline{B(0,1)} \rightarrow \overline{B(0,1)}$  admet un point fixe.

*Démonstration.* Soit une application continue  $f: \overline{B(0,1)} \rightarrow \overline{B(0,1)}$ . Nous commençons par définir une suite de fonctions

$$f_k(x) = \frac{f(x)}{1 + \frac{1}{k}}. \quad (20.111)$$

Nous avons  $\|f_k - f\|_\infty \leq \frac{1}{1+k}$  où la norme est la norme uniforme sur  $B$ . Par le théorème de Weierstrass 12.445 il existe une suite de fonctions  $C^\infty(B, \mathbb{R})$  que nous nommons  $g_k$  telles que

$$\|g_k - f_k\|_\infty \leq \frac{1}{1+k}. \quad (20.112)$$

Vérifions que cette fonction  $g_k$  soit bien une fonction qui prend ses valeurs dans  $B$  :

$$\|g_k(x)\| \leq \|g_k(x) - f_k(x)\| + \|f_k(x)\| \quad (20.113a)$$

$$\leq \frac{1}{1+k} + \frac{\|f(x)\|}{1 + \frac{1}{k}} \quad (20.113b)$$

$$\leq \frac{1}{1+k} + \frac{1}{1 + \frac{1}{k}} \quad (20.113c)$$

$$= 1. \quad (20.113d)$$

Par la version  $C^\infty$  du théorème (proposition 20.30),  $g_k$  admet un point fixe que l'on nomme  $x_k$ .

10. Notons qu'il n'existe pas non plus de rétractions continues sur  $B$ , mais pour le montrer il faut utiliser d'autres méthodes que Stokes, ou alors présenter les choses dans un autre ordre.

11. Une fonction continue sur un fermé de  $\mathbb{R}^n$  est à comprendre pour la topologie induite.

Étant donné que  $x_k$  est dans le compact  $B$ , quitte à prendre une sous-suite nous supposons que la suite  $(x_k)$  converge vers un élément  $x \in B$ . Nous montrons maintenant que  $x$  est un point fixe de  $f$  :

$$\|f(x) - x\| = \|f(x) - g_k(x) + g_k(x) - x_k + x_k - x\| \quad (20.114a)$$

$$\leq \|f(x) - g_k(x)\| + \underbrace{\|g_k(x) - x_k\|}_{=0} + \|x_k - x\| \quad (20.114b)$$

$$\leq \frac{1}{1+k} + \|x_k - x\|. \quad (20.114c)$$

En prenant le limite  $k \rightarrow \infty$  le membre de droite tend vers zéro et nous obtenons  $f(x) = x$ .  $\square$

## 20.5.2 Théorème de Schauder

Une conséquence du théorème de Brouwer est le théorème de Schauder qui est valide en dimension infinie.

ThovHJXIU

**Théorème 20.33** (Théorème de Schauder[648]).

Soit  $E$ , un espace vectoriel normé,  $K$  un convexe compact de  $E$  et  $f: K \rightarrow K$  une fonction continue. Alors  $f$  admet un point fixe.

*Démonstration.* Étant donné que  $f: K \rightarrow K$  est continue, elle y est uniformément continue. Si nous choisissons  $\epsilon$  alors il existe  $\delta > 0$  tel que

$$\|f(x) - f(y)\| \leq \epsilon \quad (20.115)$$

dès que  $\|x - y\| \leq \delta$ . La compacité de  $K$  permet de choisir un recouvrement fini par des ouverts de la forme

$$K \subset \bigcup_{1 \leq i \leq p} B(x_i, \delta) \quad (20.116)$$

EqKNPUVR

où  $\{x_1, \dots, x_p\} \subset K$ . Nous considérons maintenant  $L = \text{Span}\{f(x_j) \text{ tel que } 1 \leq j \leq p\}$  et

$$K^* = K \cap L. \quad (20.117)$$

Le fait que  $K$  et  $L$  soient convexes implique que  $K^*$  est convexe. L'ensemble  $K^*$  est également compact parce qu'il s'agit d'une partie fermée de  $K$  qui est compact (lemme 7.121). Notons en particulier que  $K^*$  est contenu dans un espace vectoriel de dimension finie, ce qui n'est pas le cas de  $K$ .

Nous allons à présent construire une sorte de partition de l'unité subordonnée au recouvrement (20.116) sur  $K$  (voir le théorème 27.88). Nous commençons par définir

$$\psi_j(x) = \begin{cases} 0 & \text{si } \|x - x_j\| \geq \delta \\ 1 - \frac{\|x - x_j\|}{\delta} & \text{sinon.} \end{cases} \quad (20.118)$$

pour chaque  $1 \leq j \leq p$ . Notons que  $\psi_j$  est une fonction positive, nulle en-dehors de  $B(x_j, \delta)$ . En particulier la fonction suivante est bien définie :

$$\varphi_j(x) = \frac{\psi_j(x)}{\sum_{k=1}^p \psi_k(x)} \quad (20.119)$$

et nous avons  $\sum_{j=1}^p \varphi_j(x) = 1$ . Les fonctions  $\varphi_j$  sont continues sur  $K$  et nous définissons finalement

$$g(x) = \sum_{j=1}^p \varphi_j(x) f(x_j). \quad (20.120)$$

Pour chaque  $x \in K$ , l'élément  $g(x)$  est une combinaison des éléments  $f(x_j) \in K^*$ . Étant donné que  $K^*$  est convexe et que la somme des coefficients  $\varphi_j(x)$  vaut un, nous avons que  $g$  prend ses valeurs dans  $K^*$  par la proposition 11.29.

Nous considérons seulement la restriction  $g: K^* \rightarrow K^*$  qui est continue sur un compact contenu dans un espace vectoriel de dimension finie. Le théorème de Brouwer 20.32 nous enseigne alors que  $g$  a un point fixe. Nous nommons  $y$  ce point fixe. Notons que  $y$  est fonction du  $\epsilon$  choisi au début de la construction, via le  $\delta$  qui avait conditionné la partition de l'unité.

Nous avons |<sup>EqoXuTzE</sup>

$$f(y) - y = f(y) - g(y) \quad (20.121a)$$

$$= \sum_{j=1}^p \varphi_j(y) f(y) - \sum_{j=1}^p \varphi_j(y) f(x_j) \quad (20.121b)$$

$$= \sum_{j=1}^p \varphi_j(y) (f(y) - f(x_j)). \quad (20.121c)$$

Par construction,  $\varphi_j(y) \neq 0$  seulement si  $\|y - x_j\| \leq \delta$  et par conséquent seulement si  $\|f(y) - f(x_j)\| \leq \epsilon$ . D'autre part nous avons  $\varphi_j(y) \geq 0$ ; en prenant la norme de (20.121) nous trouvons

$$\|f(y) - y\| \leq \sum_{j=1}^p \|\varphi_j(y) (f(y) - f(x_j))\| \leq \sum_{j=1}^p \varphi_j(y) \epsilon = \epsilon. \quad (20.122)$$

Nous nous souvenons maintenant que  $y$  était fonction de  $\epsilon$ . Soit  $y_m$  le  $y$  qui correspond à  $\epsilon = 2^{-m}$ . Nous avons alors

$$\|f(y_m) - y_m\| \leq 2^{-m}. \quad (20.123)$$

L'élément  $y_m$  est dans  $K^*$  qui est compact, donc quitte à choisir une sous-suite nous pouvons supposer que  $y_m$  est une suite qui converge vers  $y^* \in K$ <sup>12</sup>. Nous avons les majorations

$$\|f(y^*) - y^*\| \leq \|f(y^*) - f(y_m)\| + \|f(y_m) - y_m\| + \|y_m - y^*\|. \quad (20.124)$$

Si  $m$  est assez grand, les trois termes du membre de droite peuvent être rendus arbitrairement petits, d'où nous concluons que

$$f(y^*) = y^* \quad (20.125)$$

et donc que  $f$  possède un point fixe. □

### 20.5.3 Théorème de Cauchy-Arzella

**Théorème 20.34** (Cauchy-Arzela[563]).

ThoHNBooUipgPX

Nous considérons le système d'équation différentielles |<sup>EqTXlJdH</sup>

$$\begin{cases} y' = f(t, y) \end{cases} \quad (20.126a)$$

$$\begin{cases} y(t_0) = y_0. \end{cases} \quad (20.126b)$$

avec  $f: U \rightarrow \mathbb{R}^n$ , continue où  $U$  est ouvert dans  $\mathbb{R} \times \mathbb{R}^n$ . Alors il existe un voisinage fermé  $V$  de  $t_0$  sur lequel une solution  $C^1$  du problème (20.126) existe.

*Idée de la démonstration.* Nous considérons  $M = \|f\|_\infty$  et  $K$ , l'ensemble des fonctions  $M$ -Lipschitz sur  $U$ . Nous prouvons que  $(K, \|\cdot\|_\infty)$  est compact. Ensuite nous considérons l'application

$$\begin{aligned} \Phi: K &\rightarrow K \\ \Phi(f)(t) &= x_0 + \int_{t_0}^t f(u, f(u)) du. \end{aligned} \quad (20.127)$$

Après avoir prouvé que  $\Phi$  était continue, nous concluons qu'elle a un point fixe par le théorème de Schauder 20.33. □

<sup>12</sup>. Notons que même dans la sous-suite nous avons  $\|f(y_m) - y_m\| \leq 2^{-m}$ , avec le même «  $m$  » des deux côtés de l'inégalité.

**Remarque 20.35.**

Quelques remarques.

- (1) Les théorème de Cauchy-Lipschitz et Cauchy-Arzella donnent des existences pour des équations différentielles du type  $y' = f(t, y)$ . Et si nous avons une équation du second ordre ? Alors il y a la méthode de la réduction de l'ordre qui permet de transformer une équation différentielle d'ordre élevé en un système d'ordre 1.
- (2) Ces théorèmes posent des *conditions initiales* : la valeur de  $y$  est donnée en un point, et la méthode de la réduction de l'ordre permet de donner l'existence de solutions d'un problème d'ordre  $k$  en donnant les valeurs de  $y(0)$ ,  $y'(0)$ ,  $\dots y^{(k-1)}(0)$ . C'est-à-dire de la fonction et de ses dérivées en un point. Rien n'est dit sur l'existence de *conditions aux bords*.

Ces deux points sont illustrés dans les exemples 32.20 et 32.21.

**20.5.4 Théorème de Markov-Kakutani**

Le théorème de Markov-Kakutani, nous donne l'existence d'une mesure de Haar sur un groupe compact.

ThoeJCdMP

**Théorème 20.36** (Markov-Kakutani[649]).

Soit  $E$  un espace vectoriel normé et  $L$ , une partie non vide, convexe, fermée et bornée de  $E'$ . Soit  $T: L \rightarrow L$  une application linéaire et continue. Alors  $T$  a un point fixe.

*Démonstration.* Nous considérons un point  $x_0 \in L$  et la suite

$$x_n = \frac{1}{n+1} \sum_{i=0}^n T^i x_0. \quad (20.128)$$

La somme des coefficients devant les  $T^i(x_0)$  étant 1, la convexité de  $L$  montre que  $x_n \in L$ . Nous considérons l'ensemble

$$C = \bigcap_{n \in \mathbb{N}} \overline{\{x_m \text{ tel que } m \geq n\}}. \quad (20.129)$$

Le lemme 7.353 indique que  $C$  n'est pas vide, et de plus il existe une sous-suite de  $(x_n)$  qui converge vers un élément  $x \in C$ . Nous avons

$$\lim_{n \rightarrow \infty} x_{\sigma(n)}(v) = x(v) \quad (20.130)$$

pour tout  $v \in E$ . Montrons que  $x$  est un point fixe de  $T$ . Nous avons<sup>13</sup>

$$(Tx_{\sigma(k)} - x_{\sigma(k)})(v) = \frac{1}{1 + \sigma(k)} \sum_{i=0}^{\sigma(k)} (T^{i+1}x_0)(v) - \frac{1}{1 + \sigma(k)} \sum_{i=0}^{\sigma(k)} (T^i x_0)(v) \quad (20.131a)$$

$$= \frac{1}{1 + \sigma(k)} \sum_{i=0}^{\sigma(k)} (T^{i+1}x_0 - T^i x_0)(v) \quad (20.131b)$$

$$= \frac{1}{1 + \sigma(k)} (T^{i+\sigma(k)+1}x_0 - x_0)(v) \quad \text{somme télescopique.} \quad (20.131c)$$

En prenant la norme,

$$\|(Tx_{\sigma(k)} - x_{\sigma(k)})v\| = \frac{1}{1 + \sigma(k)} \|T^{\sigma(k)+1}x_0(v) - x_0(v)\| \quad (20.132a)$$

$$\leq \frac{2M}{\sigma(k) + 1} \quad (20.132b)$$

où  $M = \sum_{y \in L} \|y(v)\| < \infty$  parce que  $L$  est borné. En prenant  $k \rightarrow \infty$  nous trouvons

$$\lim_{k \rightarrow \infty} (Tx_{\sigma(k)} - x_{\sigma(k)})v = 0, \quad (20.133)$$

ce qui signifie que  $Tx = x$  parce que  $T$  est continue. □

13. Je ne vois pas comment fonctionne ce passage si on suppose que  $T$  est affine et non linéaire.

## 20.6 Intégrales curvilignes

### 20.6.1 Chemins de classe $C^1$

secintcurvi

DEFooQTAWooOCmSJo

**Définition 20.37** ([650]).

Soit  $p, q \in \mathbb{R}^n$ . Un **chemin**  $C^1$  **par morceaux** joignant  $p$  à  $q$  est une application continue

$$\gamma : [a, b] \rightarrow \mathbb{R}^n \quad (20.134)$$

vérifiant :

- (1) L'application  $\gamma : [a, b] \rightarrow \mathbb{R}^n$  est continue.
- (2) Les extrémités :  $\gamma(a) = p$ ,  $\gamma(b) = q$
- (3) Il existe une subdivision  $a = t_0 < t_1 < \dots < t_{r-1} < t_r = b$  telle que la restriction de  $\gamma$  sur chaque fermé  $[t_i, t_{i+1}]$  est de classe  $C^1$ .

Pour être d'accord sur les définitions, quand on dit qu'une application est  $C^1$  sur un fermé, c'est qu'elle est  $C^1$  sur ce fermé, avec la topologie de ce fermé (induite<sup>14</sup> de  $\mathbb{R}$ ). Autrement dit, toutes les limites et tous les ouverts qui interviennent dans la notion de « être  $C^1$  » sont à prendre dans la topologie du fermé  $[a_i, b_i]$ .

**Remarque 20.38.**

Si  $a$  et  $b$  sont des points de  $\mathbb{R}^n$ , on peut créer le chemin particulier

$$\gamma : [0, 1] \rightarrow \mathbb{R}^n : t \mapsto (1 - t)a + tb \quad (20.135)$$

qui relie ces points par un segment de droite.

Il existe des fonctions  $C^1$  par morceaux qui ne sont pas  $C^1$ .

**Proposition 20.39.**

L'application

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \begin{cases} x & \text{si } x \leq 0 \\ 2x & \text{si } x > 0 \end{cases} \end{aligned} \quad (20.136)$$

est  $C^1$  par morceaux.

Elle n'est pas de classe  $C^1$ .

*Démonstration.* Le fait que  $f$  ne soit pas de classe  $C^1$  est immédiat parce que sa dérivée n'est pas définie en zéro.

Pour montrer que  $f$  est  $C^1$  par morceaux, nous définissons les deux fonctions

$$\begin{aligned} f_1 : ]-\infty, 0] &\rightarrow \mathbb{R} \\ x &\mapsto x \end{aligned} \quad (20.137)$$

et

$$\begin{aligned} f_2 : [0, \infty[ &\rightarrow \mathbb{R} \\ x &\mapsto 2x. \end{aligned} \quad (20.138)$$

Nous calculons la dérivée de  $f_1$  en  $x = 0$ . C'est par définition

$$\lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x} = \lim_{x \rightarrow 0} \frac{f(x)}{x}. \quad (20.139)$$

Attention : cette limite est à comprendre au sens de la topologie de  $]-\infty, 0]$ . Nous allons montrer que cette limite vaut 1, en remontant aux définition pour le sport.

14. Topologie induite, définition 7.27.

Posons  $g(x) = f(x)/x$ . Si  $V$  est un voisinage de 1 dans  $\mathbb{R}$ , il existe une boule  $B(1, r)$  qui contient  $V$ . Nous allons trouver un  $\epsilon > 0$  tel que  $g(B(0, \epsilon) \setminus \{0\}) \subset B(1, r)$ . Attention : à gauche c'est une boule au sens de la topologie induite. Au sens de la topologie normale, nous cherchons  $\epsilon > 0$  tel que

$$g\left((B(0, \epsilon) \cap ]-\infty, 0]) \setminus \{0\}\right) \quad (20.140)$$

Bref, pour tout  $x$  dans cette boule, nous devons avoir  $|g(x) - 1| < r$ . Vu que  $x \in ]-\infty, 0] \setminus \{0\}$ , nous avons  $g(x) = 1$  pour tout  $x$ . Et voilà, n'importe que  $\epsilon$  fait l'affaire.

Nous avons prouvé que la dérivée de  $f_1$  est donnée par

$$\begin{aligned} f'_1 : ]-\infty, 0] &\rightarrow \mathbb{R} \\ x &\mapsto 1. \end{aligned} \quad (20.141)$$

Elle est continue. Nous laissons la démonstration de la continuité de  $f'_1$  à la lectrice ; attention : continuité pour la topologie induite, toujours. Bref,  $f_1$  est de classe  $C^1$ .

Nous démontrons de même que  $f_2$  est de classe  $C^2$  et que  $f'_2(x) = 2$  pour tout  $x$ .  $\square$

### 20.6.2 Intégrer une fonction

DEFooFAYUooCaUdyo

#### Définition 20.40.

Soit  $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue, et  $\gamma : [a, b] \rightarrow D$  un chemin  $C^1$ . On définit l'intégrale de  $f$  sur  $\gamma$  par

$$\int_{\gamma} f ds = \int_{\gamma} f = \int_a^b f(\gamma(t)) \|\gamma'(t)\| dt. \quad (20.142) \quad \text{EghJGRcb}$$

Note : dans le cadre de l'analyse complexe, ce n'est pas exactement cette définition. Voir 26.6.

#### Exemple 20.41.

Soit l'hélice

$$\begin{aligned} \sigma : [0, 2\pi] &\rightarrow \mathbb{R}^3 \\ t &\mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix}, \end{aligned} \quad (20.143)$$

et la fonction  $f(x, y, z) = x^2 + y^2 + z^2$ . L'intégrale de  $f$  sur  $\sigma$  est

$$\begin{aligned} \int_{\sigma} f &= \int_0^{2\pi} (\cos^2 t + \sin^2 t + t^2) \|\sigma'(t)\| dt \\ &= \int_0^{2\pi} (1 + t^2) \sqrt{2} dt \\ &= \sqrt{2} \left[ t + \frac{t^3}{3} \right]_0^{2\pi} \\ &= \sqrt{2} \left( 2\pi + \frac{8\pi^3}{3} \right). \end{aligned} \quad (20.144)$$

$\triangle$

#### Remarque 20.42.

Si  $f = 1$ , alors nous tombons sur

$$\int_{\gamma} ds = \int_a^b \|\gamma'(t)\| dt, \quad (20.145)$$

Nous verrons par le théorème 21.9 que cette dernière intégrale est la longueur de la courbe. Il est un fait général que l'intégrale de la fonction 1 sur un ensemble en donne la « mesure ». Cela est à mettre en rapport avec le lemme 14.175 en gardant en tête que  $\int_{\gamma} 1$  n'est pas la mesure de l'image de  $\gamma$  dans  $\mathbb{R}^2$ .



**Proposition 20.43** (Indépendance en le paramétrage).

La valeur de l'intégrale de  $f$  sur  $\gamma$  ne dépend pas du paramétrage (équivalent ou pas) choisi.

*Démonstration.* Soit donc un chemin  $\gamma: [c, d] \rightarrow \mathbb{R}^3$  ainsi que  $\varphi: [c, d] \rightarrow [a, b]$ , un reparamétrage de classe  $C^1$ , strictement monotone et le chemin  $\sigma$  défini par  $\gamma(s) = \sigma(\varphi(s))$  avec  $s \in [c, d]$ . En supposant que  $\varphi'(s) \geq 0$ , nous avons

$$\begin{aligned} I &= \int_{\gamma} f = \int_c^d f(\gamma(s)) \|\gamma'(s)\| ds \\ &= \int_c^d f(\sigma(\varphi(s))) \|\sigma'(\varphi(s))\| |\varphi'(s)| ds. \end{aligned} \quad (20.146)$$

Pour cette intégrale, nous posons  $t = \varphi(s)$ , et par conséquent  $dt = \varphi'(s)ds$ . Étant donné que  $\varphi'(s) \geq 0$ , nous pouvons supprimer les valeurs absolues, et obtenir

$$\begin{aligned} I &= \int_{\varphi(c)}^{\varphi(d)} f(\sigma(t)) \|\sigma'(t)\| dt \\ &= \int_a^b f(\sigma(t)) \|\sigma'(t)\| dt \\ &= \int_{\sigma} f. \end{aligned} \quad (20.147)$$

Essayez de faire le cas  $\varphi'(s) \leq 0$ . □

RemiqswPd

**Remarque 20.44.**

Attention : les intégrales sur des chemins dans  $\mathbb{C}$  ne sont la même chose. En effet  $\mathbb{C}$  doit être souvent plutôt traité comme  $\mathbb{R}$  que comme  $\mathbb{R}^2$ . Si  $\gamma$  est un chemin dans  $\mathbb{C}$ , l'intégrale

$$\int_{\gamma} f \quad (20.148)$$

doit être comprise comme une généralisation de  $\int_a^b f(x)dx$  et non comme l'intégrale sur un chemin. La différence est qu'en retournant les bornes d'une intégrale usuelle sur  $\mathbb{R}$  on change le signe, alors qu'en retournant un chemin dans  $\mathbb{R}^2$ , on ne change pas. Bref, la définition est que si  $\gamma: [a, b] \rightarrow \mathbb{C}$  est un chemin, alors

$$\int_{\gamma} f = \int_{\gamma} f(z)dz = \int_a^b f(\gamma(t))\gamma'(t)dt, \quad (20.149)$$

sans valeur absolue autour de  $\gamma'(t)$ .

### 20.6.3 Intégrer un champ de vecteurs

DEFooSHHFooVdsxMf

**Définition 20.45.**

Un **champ de vecteur** est une application  $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ . On définit l'intégrale de  $G$  sur un chemin  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  par

$$\int_{\gamma} G \stackrel{\text{def}}{=} \int_a^b \langle G(\gamma(t)), \gamma'(t) \rangle dt.$$

**Remarque 20.46.**

Cette définition ne dépend pas du paramétrage choisi, mais le signe change selon le sens du chemin.

### 20.6.4 Intégrer une forme différentielle sur un chemin

La formule d'intégration d'un champ de vecteur <sup>15</sup>,

$$\int_{\gamma} G = \int_{[a,b]} \langle G(\gamma(t)), \gamma'(t) \rangle dt, \quad (20.150)$$

---

15. Définition 20.45.

contient quelque chose d'intéressant : la combinaison  $\langle G(\gamma(t)), \gamma'(t) \rangle$ . Cette combinaison sert à transformer le vecteur tangent  $\gamma'(t)$  en un nombre en utilisant le produit scalaire avec le vecteur  $G(\gamma(t))$ .

Si  $G$  est un champ de vecteur sur  $\mathbb{R}^n$ , et si  $x \in \mathbb{R}^n$ , nous pouvons utiliser l'isomorphisme musical (définition 12.676)

$$\begin{aligned} G_x^b: \mathbb{R}^n &\rightarrow \mathbb{R} \\ v &\mapsto \langle G(x), v \rangle \end{aligned} \quad (20.151)$$

pour écrire de façon plus compacte :

$$\int_{\gamma} G = \int_{[a,b]} G_{\gamma(t)}^b(\gamma'(t)) dt. \quad (20.152)$$

#### Définition 20.47.

Soient une forme différentielle  $\omega$  sur  $\mathbb{R}^n$  et un chemin de classe  $C^1$   $\gamma: [a, b] \rightarrow \mathbb{R}^n$ . L'intégrale de  $\omega$  sur  $\gamma$  est définie par

$$\int_{\gamma} \omega = \int_a^b \omega_{\gamma(t)} \gamma'(t) dt \quad (20.153)$$

#### Remarque 20.48.

Cette définition ne dépend pas du paramétrage choisi, mais le signe change selon le sens du chemin.

### 20.6.5 Intégration d'une forme différentielle sur un chemin

Les formes intégrales que nous avons déjà vues sont celles de fonctions et de champs de vecteur sur des chemins. Si  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  est le chemin, les formules sont

$$\begin{aligned} \int_{\gamma} f &= \int_{[a,b]} f(\gamma(t)) \|\gamma'(t)\| dt \\ \int_{\gamma} G &= \int_{[a,b]} \langle G(\gamma(t)), \gamma'(t) \rangle dt. \end{aligned} \quad (20.154)$$

Dans les deux cas, le principe est que nous disposons de quelque chose (la fonction  $f$  ou le vecteur  $G$ ), et du vecteur tangent  $\gamma'(t)$ , et nous essayons d'en tirer un nombre que nous intégrons. Lorsque nous avons une 1-forme, la façon de l'utiliser pour produire un nombre avec le vecteur tangent est évidemment d'appliquer la 1-forme au vecteur tangent. La définition suivante est donc naturelle.

#### Définition 20.49.

Soit  $\gamma: [a, b] \rightarrow \mathbb{R}^n$ , un chemin de classe  $C^1$  tel que son image est contenue dans le domaine  $D$ . Si  $\omega$  est une 1-forme différentielle sur  $D$ , nous définissons l'intégrale de  $\omega$  le long de  $\gamma$  le nombre

$$\begin{aligned} \int_{\gamma} \omega &= \int_a^b \omega_{\gamma(t)}(\gamma'(t)) dt \\ &= \int_a^b [a_1(\gamma(t))\gamma'_1(t) + \cdots + a_n(\gamma(t))\gamma'_n(t)] dt. \end{aligned} \quad (20.155)$$

Cette définition est une bonne définition parce que si on change le paramétrage du chemin, on ne change pas la valeur de l'intégrale, c'est la proposition suivante.

#### Proposition 20.50.

Si  $\gamma$  et  $\beta$  sont des chemins équivalents, alors

$$\int_{\gamma} \omega = \int_{\beta} \omega, \quad (20.156)$$

c'est-à-dire que l'intégrale est invariante sous les reparamétrages du chemin.

*Démonstration.* Deux chemins sont équivalents quand il existe un difféomorphisme  $C^1$   $h: [a, b] \rightarrow [c, d]$  tel que  $\gamma(t) = (\beta \circ h)(t)$ . En remplaçant  $\gamma$  par  $(\beta \circ h)$  dans la définition de  $\int_\gamma \omega$ , nous trouvons

$$\int_a^b \omega_{\gamma(t)}(\gamma'(t)) dt = \int_a^b \omega_{(\beta \circ h)(t)}((\beta \circ h)'(t)) dt. \quad (20.157)$$

Un changement de variable  $u = h(t)$  transforme cette dernière intégrale en  $\int_\beta \omega$ , ce qui prouve la proposition.  $\square$

**Remarque 20.51.**

Si  $\gamma$  est une somme de chemins,  $\gamma = \gamma^{(1)} + \dots + \gamma^{(n)}$ , où chacun des  $\gamma^{(i)}$  est un chemin, alors

$$\int_\gamma \omega = \sum_{i=1}^n \int_{\gamma_i} \omega \quad (20.158)$$

parce que  $\omega$  est linéaire.

**Remarque 20.52.**

Si  $-\gamma$  est le chemin

$$\begin{aligned} -\gamma: [a, b] &\rightarrow \mathbb{R}^n \\ t &\mapsto \gamma(b - (t - a)), \end{aligned} \quad (20.159)$$

alors

$$\int_{-\gamma} \omega = - \int_\gamma \omega, \quad (20.160)$$

c'est-à-dire que si l'on parcourt le chemin en sens inverse, alors on change le signe de l'intégrale.

L'intégrale d'une forme différentielle sur un chemin est compatible avec l'intégrale déjà connue d'un champ de vecteur sur le chemin parce que si  $G$  est un champ de vecteurs,

$$\int_\gamma G^b = \int_\gamma G. \quad (20.161)$$

En effet,

$$\begin{aligned} \int_\gamma G^b &= \int_a^b G_{\gamma(t)}^b(\gamma'(t)) \\ &= \int_a^b [G_1(\gamma(t))dx_1 + \dots + G_n(\gamma(t))dx_n](\gamma'_1(t), \dots, \gamma'_n(t)) \\ &= \int_a^b \langle G(\gamma(t)), \gamma'(t) \rangle \\ &= \int_\gamma G. \end{aligned} \quad (20.162)$$

**Proposition 20.53.**

Soit  $\omega = df$ , une 1-forme exacte et continue sur le domaine  $D$ . Alors la valeur de  $\int_\gamma df$  ne dépend que des valeurs de  $f$  aux extrémités de  $\gamma$ .

*Démonstration.* Nous avons

$$\begin{aligned} \int_\gamma \omega &= \int_\gamma df = \int_a^b \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t)) \gamma'_i(t) dt \\ &= \int_a^b \frac{d}{dt} ((f \circ \gamma)(t)) dt \\ &= (f \circ \gamma)(b) - (f \circ \gamma)(a). \end{aligned} \quad (20.163)$$

$\square$

### 20.6.6 Interprétation physique : travail

**Définition 20.54** ([651]).

Une force  $F: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  est **conservative** si elle dérive d'un potentiel, c'est-à-dire si il existe une fonction  $V \in C^1(D, \mathbb{R})$  telle que

$$F(x) = (\nabla V)(x). \quad (20.164)$$

Étant donné que  $F$  est un champ de vecteurs, nous avons une forme différentielle associée  $F^\flat$ ,

$$F_x^\flat: v \mapsto \langle F(x), v \rangle. \quad (20.165)$$

**Lemme 20.55.**

Le champ  $F$  est conservatif si et seulement si la 1-forme différentielle  $F^\flat$  est exacte.

*Démonstration.* Supposons que la force  $F$  soit conservative, c'est-à-dire qu'il existe une fonction  $V$  telle que  $F = \nabla V$ . Dans ce cas, il est facile de prouver que  $F^\flat$  est exacte et est donnée par  $F_x^\flat = dV(x)$ . En effet,

$$\begin{aligned} F_x^\flat(v) &= \langle F(x), v \rangle \\ &= F_1(x)v_1 + \dots + F_n(x)v_n \\ &= \frac{\partial V}{\partial x_1}(x)v_1 + \dots + \frac{\partial V}{\partial x_n}(x)v_n \\ &= dV(x)v. \end{aligned} \quad (20.166)$$

Pour le sens inverse, supposons que  $F^\flat$  soit exacte. Dans ce cas, nous avons une fonction  $V$  telle que  $F^\flat = dV$ . Il est facile de prouver qu'alors,  $F = \nabla V$ .  $\square$

En résumé, nous avons deux façons équivalentes d'exprimer que la force  $F$  dérive du potentiel  $V$  : soit nous disons  $F = \nabla V$ , soit nous disons  $F^\flat = dV$ .

**Proposition 20.56.**

Si  $F$  est une force conservative, alors le travail<sup>16</sup> de  $F$  lors d'un déplacement ne dépend pas du chemin suivi.

*Démonstration.* Le travail d'une force le long d'un chemin n'est autre que l'intégrale de la force le long du chemin, et le calcul est facile :

$$W_\gamma(F) = \int_\gamma F = \int_\gamma dV = V(\gamma(b)) - V(\gamma(a)). \quad (20.167)$$

Donc si  $\beta$  est un autre chemin tel que  $\beta(a) = \gamma(a)$  et  $\beta(b) = \gamma(b)$ , nous avons  $W_\beta(F) = W_\gamma(F)$ .  $\square$

### 20.6.7 Intégrer un champ de vecteurs sur un bord en $2D$

Si  $D \subset \mathbb{R}^2$  est tel que  $\partial D$  est une variété de dimension 1 et tel que  $D$  accepte un champ de vecteur normal extérieur unitaire  $\nu$ . Si nous voulons définir

$$\int_{\partial D} G, \quad (20.168)$$

le mieux est de prendre un paramétrage  $\gamma: [0, 1] \rightarrow \mathbb{R}^2$  et de calculer

$$\int_0^1 \langle G_{\gamma(t)}, \frac{\dot{\gamma}(t)}{\|\dot{\gamma}(t)\|} \rangle dt. \quad (20.169)$$

Hélas, cette définition ne fonctionne pas parce que son signe dépend du sens du paramétrage  $\gamma$ . Si le paramétrage tourne dans l'autre sens, il y a un signe de différence.

16. Voir [652].

Nous allons définir

$$\int_{\partial D} G = \int_0^1 \langle G_{\gamma(t)}, T(t) \rangle dt \quad (20.170) \quad \text{EqIntVectbordDeux}$$

où  $T(t) = \dot{\gamma}(t)/\|\dot{\gamma}(t)\|$  et où  $\gamma$  est choisi de telle façon que la rotation d'angle  $\frac{\pi}{2}$  amène  $\nu$  sur  $T$ . Cela fixe le choix de sens.

Ce choix de sens aura des répercussions dans l'application de la formule de Green et du théorème de Stokes.

### 20.6.8 Intégrer une forme différentielle sur un bord en $2D$

Nous n'allons pas chercher très loin :

$$\int_{\partial D} \omega = \int_{\partial D} \omega^\sharp, \quad (20.171)$$

c'est-à-dire que l'intégrale de la forme différentielle est celle du champ de vecteur associé. Le membre de droite est défini par (20.170), avec le choix d'orientation qui va avec.

### 20.6.9 Intégrer une forme différentielle sur un bord en $3D$

Nous allons maintenant intégrer une forme différentielle sur certains chemins fermés dans  $\mathbb{R}^3$ . Soit  $F(D) \subset \mathbb{R}^3$ , une variété de dimension 2 dans  $\mathbb{R}^3$  où  $F: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  est la carte. Nous supposons que  $D$  vérifie les hypothèses de la formule de Green. Alors nous définissons

$$\int_{F(\partial D)} \omega = \int_{\partial D} F^* \omega \quad (20.172) \quad \text{EqDefIntTroisForBord}$$

où  $F^* \omega$  est la forme différentielle définie sur  $\partial D$  par  $(F^* \omega)(v) = \omega(dF(v))$ .

Cette définition est très abstraite, mais nous n'allons, en pratique, jamais l'utiliser, grâce au théorème de Stokes.

### 20.6.10 Intégrer un champ de vecteurs sur un bord en $3D$

Encore une fois, nous n'allons pas chercher bien loin :

$$\int_{F(\partial D)} G = \int_{F(\partial D)} G^\flat \quad (20.173)$$

où  $G^\flat$  est la forme différentielle associée au champ de vecteur. Le membre de droite est défini par l'équation (20.172).

### 20.6.11 Dérivées croisées et forme différentielle exacte

Nous considérons le problème suivant : trouver une fonction  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  telle que l'EqskfgfNr

$$\begin{cases} \frac{\partial f}{\partial x} = a(x, y) \\ \frac{\partial f}{\partial y} = b(x, y) \end{cases} \quad (20.174a) \quad (20.174b)$$

où  $a$  et  $b$  sont des fonctions supposées suffisamment régulières. Nous savons que ce problème n'a pas de solutions lorsque

$$\frac{\partial a}{\partial y} \neq \frac{\partial b}{\partial x} \quad (20.175)$$

parce que cela impliquerait  $\partial_{xy}^2 f \neq \partial_{yx}^2 f$ . Nous sommes en droit de nous demander si la condition

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x} \quad (20.176)$$

impliquerait qu'il existe une solution au problème (20.174). La réponse est oui, et nous allons brièvement la justifier. Pour plus de détails nous vous demandons de chercher un peu. La référence [653] peut être utile.

**Proposition 20.57.**

Si  $a$  et  $b$  sont des fonctions qui satisfont à la condition

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}, \quad (20.177)$$

alors la fonction

$$f(x, y) = \int_0^x a(t, 0)dt + \int_0^y b(x, t)dt \quad (20.178)$$

répond au problème

$$\begin{cases} \frac{\partial f}{\partial x} = a(x, y) \\ \frac{\partial f}{\partial y} = b(x, y) \end{cases} \quad (20.179a)$$

$$(20.179b)$$

La preuve qui suit n'en est pas complètement une parce qu'il manque des justifications, notamment au moment de permuter la dérivée et l'intégrale.

*Démonstration.* La clef de la preuve est le théorème fondamental de l'analyse :

$$\int_0^x \frac{\partial f}{\partial x}(t, y)dt = f(x, y) \quad (20.180)$$

et son pendant par rapport à  $y$  :

$$\int_0^y \frac{\partial f}{\partial y}(x, t)dt = f(x, y). \quad (20.181)$$

En appliquant ces version du théorème fondamental, nous obtenons immédiatement.

$$\frac{\partial f}{\partial y} = b(x, y). \quad (20.182)$$

En ce qui concerne la dérivée par rapport à  $y$ ,

$$\frac{\partial f}{\partial x} = a(x, 0) + \int_0^y \frac{\partial b}{\partial x}(x, t)dt \quad (20.183a)$$

$$= a(x, 0) + \int_0^y \frac{\partial a}{\partial y}(x, t)dt \quad (20.183b)$$

$$= a(x, 0) + [a(x, t)]_{t=0}^{t=y} \quad (20.183c)$$

$$= a(x, y). \quad (20.183d)$$

□

En ce qui concerne l'unicité, supposons que  $f$  et  $g$  soient deux solutions au problème. L'équation

$$\frac{\partial f}{\partial x} = a(x, y) = \frac{\partial g}{\partial x} \quad (20.184)$$

implique que

$$f(x, y) = g(x, y) + C(y) \quad (20.185)$$

où  $C$  est une fonction seulement de  $y$ . L'autre équation implique

$$f(x, y) = g(x, y) + D(x) \quad (20.186)$$

où  $D$  est seulement une fonction de  $x$ . En égalisant nous voyons que les fonctions  $C$  et  $D$  doivent être des constantes.

Par conséquent la fonction  $f$  est donnée à une constante près et en réalité la fonction (20.178) est suffisante pour répondre au problème de trouver toutes les fonctions dont les dérivées partielles sont données par les fonctions  $a$  et  $b$ .

La fonction  $f$  ainsi créée est un **potentiel** pour le champ de force

$$F(x, y) = \begin{pmatrix} a(x, y) \\ b(x, y) \end{pmatrix}. \quad (20.187)$$

Notez que ce champ de vecteurs est le gradient de  $f$ . La question initiale aurait donc pu être posée en les termes suivants : trouver une fonction  $f$  dont le gradient est donné par

$$\nabla f = \begin{pmatrix} a(x, y) \\ b(x, y) \end{pmatrix}. \quad (20.188)$$

## 20.7 Surfaces paramétrées

De la même façon qu'un chemin dans  $\mathbb{R}^3$  est décrit comme une application  $\sigma: \mathbb{R} \rightarrow \mathbb{R}^3$ , une surface dans  $\mathbb{R}^3$  sera vue comme une application  $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ . Une **surface paramétrée** dans  $\mathbb{R}^3$  est une application

$$\begin{aligned} \varphi: D \subset \mathbb{R}^2 &\rightarrow \mathbb{R}^3 \\ (u, v) &\mapsto \varphi(u, v) = \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{pmatrix}. \end{aligned} \quad (20.189)$$

Nous allons parler de la « surface  $\varphi$  » pour désigner l'image de  $\varphi$  dans  $\mathbb{R}^3$ .

Si on fixe le paramètre  $u = u_0$ , alors l'application

$$v \mapsto \varphi(u_0, v) \quad (20.190)$$

est un chemin dans la surface. Un vecteur tangent à ce chemin sera tangent à la courbe :

$$\frac{\partial \varphi}{\partial v}(u_0, v_0) = \begin{pmatrix} \frac{\partial x}{\partial v}(u_0, v_0) \\ \frac{\partial y}{\partial v}(u_0, v_0) \\ \frac{\partial z}{\partial v}(u_0, v_0) \end{pmatrix}. \quad (20.191)$$

De même, en fixant  $v_0$ , on considère le chemin

$$u \mapsto \varphi(u, v_0). \quad (20.192)$$

Le vecteur tangent à ce chemin est égal à tangent à la surface :

$$\frac{\partial \varphi}{\partial u} = \begin{pmatrix} \frac{\partial x}{\partial u}(u_0, v_0) \\ \frac{\partial y}{\partial u}(u_0, v_0) \\ \frac{\partial z}{\partial u}(u_0, v_0) \end{pmatrix} \quad (20.193)$$

### Définition 20.58.

Nous disons que la surface est **régulière** si les vecteurs  $\partial_u \varphi(u_0, v_0)$  et  $\partial_v \varphi(u_0, v_0)$  sont non nuls et non colinéaires.

Si la surface est régulière, les vecteurs tangents au paramétrage forment le plan tangent à la surface au point  $\varphi(u_0, v_0)$ .

Un vecteur orthogonal à la surface (et donc au plan tangent) est donc donné par le produit vectoriel :

$$n(u_0, v_0) = \frac{\partial \varphi}{\partial u}(u_0, v_0) \times \frac{\partial \varphi}{\partial v}(u_0, v_0). \quad (20.194)$$

DefSurfReguliere

L'équation du plan tangent est alors obtenue par

$$\begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \cdot n(u_0, v_0) = 0 \quad (20.195)$$

où  $x_0 = x(u_0, v_0)$ ,  $y_0 = y(u_0, v_0)$ ,  $z_0 = z(u_0, v_0)$ .

### 20.7.1 Graphe d'une fonction

Soit la fonction  $f: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ . Le graphe de  $f$  est l'ensemble des points de la forme

$$(x, y, f(x, y)) \quad (20.196)$$

tels que  $(x, y) \in D$ . Cela est une surface paramétrée par

$$\begin{aligned} \varphi: D &\rightarrow \mathbb{R}^3 \\ (x, y) &\mapsto \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix}. \end{aligned} \quad (20.197)$$

Les vecteurs tangents sont

$$\frac{\partial \varphi}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x} \end{pmatrix}, \quad \frac{\partial \varphi}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y} \end{pmatrix}. \quad (20.198)$$

La surface est donc partout régulière parce que ces deux vecteurs ne sont jamais nuls ou colinéaires. Un vecteur normal à cette surface au point  $(x_0, y_0, f(x_0, y_0))$  est donné par le produit vectoriel

$$n = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & \frac{\partial f}{\partial x}(x_0, y_0) \\ 0 & 1 & \frac{\partial f}{\partial y}(x_0, y_0) \end{vmatrix} = -\frac{\partial f}{\partial x}(x_0, y_0)e_x - \frac{\partial f}{\partial y}(x_0, y_0)e_y + e_z. \quad (20.199)$$

En suivant l'équation (20.195), nous avons l'équation suivante pour le plan :

$$\begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \cdot \begin{pmatrix} -\frac{\partial f}{\partial x}(x_0, y_0) \\ -\frac{\partial f}{\partial y}(x_0, y_0) \\ 1 \end{pmatrix} = 0, \quad (20.200)$$

c'est-à-dire

$$-(x - x_0)\frac{\partial f}{\partial x}(x_0, y_0) - (y - y_0)\frac{\partial f}{\partial y}(x_0, y_0) + z - f(x_0, y_0) = 0, \quad (20.201)$$

ce qui revient à

$$z - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0). \quad (20.202)$$

Nous retrouvons donc l'équation du plan tangent à un graphe.

#### Exemple 20.59.

La sphère de rayon  $R$  peut être paramétrée par les angles sphériques :

$$\phi(\theta, \varphi) = \begin{pmatrix} R \sin \theta \cos \varphi \\ R \sin \theta \sin \varphi \\ R \cos \theta \end{pmatrix} \quad (20.203)$$

avec  $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi]$ .



Tentons d'en trouver le plan tangent au point  $(x, y, z) = (R, 0, 0)$ . Un petit dessin nous montre que c'est un plan vertical d'équation  $x = R$ . Montrons cela en utilisant la théorie que nous venons de découvrir. D'abord le point  $(R, 0, 0)$  correspond à  $\theta_0 = \frac{\pi}{2}$  et  $\varphi = 0$ . Les vecteurs tangents sont

$$T_\theta = \frac{\partial \phi}{\partial \theta}(R, \frac{\pi}{2}, 0) = \begin{pmatrix} R \cos \theta \cos \varphi \\ R \cos \theta \sin \varphi \\ -R \sin \theta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -R \end{pmatrix}, \quad \text{EqTthetaSph} \quad (20.204)$$

et

$$T_\varphi = \frac{\partial \phi}{\partial \varphi}(R, \frac{\pi}{2}, 0) = \begin{pmatrix} -R \sin \theta \sin \varphi \\ R \sin \theta \cos \varphi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ R \\ 0 \end{pmatrix}. \quad \text{EqTvarphiSph} \quad (20.205)$$

Cela sont de toute évidence bien les deux vecteurs tangents à la sphère au point  $(x, y, z) = (R, 0, 0)$ . Le vecteur normal est

$$\begin{vmatrix} e_x & e_y & e_z \\ 0 & 0 & -R \\ 0 & R & 0 \end{vmatrix} = R^2 e_x. \quad (20.206)$$

Ici encore, nous avons le vecteur que nous attendions sur un dessin. L'équation du plan tangent est maintenant

$$\begin{pmatrix} x - R \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} R^2 \\ 0 \\ 0 \end{pmatrix} = 0, \quad (20.207)$$

c'est-à-dire  $R^2(x - R) = 0$  et donc  $x = R$ . △

### 20.7.2 Intégrale sur une partie de $\mathbb{R}^m$

Soit  $M$  une variété de dimension  $n$  dans  $\mathbb{R}^m$ . Soit  $F : W \subset \mathbb{R}^n \rightarrow M$  un paramétrage d'un ouvert relatif de  $M$ .

Si  $f$  est une fonction définie sur un sous-ensemble  $A \subset F(W)$  tel que  $F^{-1}(A)$  est mesurable, l'intégrale de  $f$  sur  $A$  est définie par

$$\int_A f = \int_{F^{-1}(A)} f(F(w)) \sqrt{\det(J_F(w)^t J_F(w))} dw$$

où l'intégrale est l'intégration usuelle (de Lebesgue) sur  $F^{-1}(A) \subset \mathbb{R}^n$ . On écrit parfois cette intégrale  $\int_{F^{-1}(A)} f(F(w)) d\sigma$  où

$$d\sigma = \sqrt{\det(J_F(w)^t J_F(w))} dw \quad \text{EqQoARMAooQPhQAL} \quad (20.208)$$

est l'élément infinitésimal de volume de la variété.

Si  $m = 3$  et  $n = 2$ , l'élément infinitésimal de volume vaut

$$d\sigma = \left\| \frac{\partial F}{\partial w_1} \times \frac{\partial F}{\partial w_2} \right\| dw$$

où  $\times$  représente le produit vectoriel dans  $\mathbb{R}^3$ , et  $(w_1, w_2)$  sont les coordonnées sur  $W \subset \mathbb{R}^2$ . Dans la suite, nous ne regarderons plus que ce cas.

## 20.8 Intégrales de surface

### 20.8.1 Intégrale d'un champ de vecteurs

Dans l'intégration curviligne, on a noté que si l'intégrale d'une fonction ne dépendait pas de l'orientation du chemin, l'intégrale d'un champ de vecteurs ou d'une forme différentielle en dépendait. Ce problème d'orientation apparaît également dans l'intégration sur des surfaces de l'espace.

secintsurfaciques

DEFooFTQLooXXbtOQ

**Définition 20.60.**

Une **orientation** sur une surface  $S \subset \mathbb{R}^3$  est le choix d'un champ de vecteurs continu  $\nu : S \rightarrow \mathbb{R}^3$  dont la norme en tout point de  $S$  vaut 1.

On remarque qu'ayant fait un tel choix d'orientation  $\nu(x)$  en un point  $x$ , le seul autre choix possible en  $x$  est  $-\nu(x)$ . Si  $S$  est le bord d'un ouvert  $D \subset \mathbb{R}^3$ , l'*orientation induite par  $D$  sur  $S$*  est, si elle existe, l'orientation qui pointe hors de  $D$  en tout point de  $S$ . Plus précisément, il faut que pour tout  $x \in D$  il existe  $\epsilon > 0$  vérifiant, pour tout  $0 < t < \epsilon$ , la relation  $t\nu(x) \notin D$ . Dans ce cas, le champ de vecteurs  $\nu$  est appelé le *vecteur normal unitaire extérieur* à  $D$  et il est forcément unique.

Soit  $G$  un champ de vecteurs défini sur une surface orientée par un champ  $\nu$ . L'intégrale de  $G$  sur  $S$ , aussi appelée le *flux de  $G$  à travers  $S$* , est

$$\iint_S G \cdot dS \stackrel{\text{def}}{=} \iint_S \langle G, \nu \rangle d\sigma. \quad (20.209) \quad \text{sgflux-star}$$

Si on suppose que la surface est paramétrée par une application

$$F : W \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3 : (u, v) \mapsto (F_1(u, v), F_2(u, v), F_3(u, v))$$

alors un vecteur unitaire  $\nu$  peut s'écrire sous la forme

$$\nu = \frac{\frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v}}{\left\| \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\|}$$

et grâce à ce paramétrage l'intégrale (20.209) devient

$$\iint_S G \cdot dS = \iint_W \left\langle G(F(u, v)), \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\rangle du dv.$$

où on utilise l'expression de  $d\sigma$  obtenue précédemment dans le cas qui nous intéresse (surface dans l'espace).

## 20.9 Aires et intégrales

### 20.9.1 Aire d'une surface paramétrée

Lorsque nous avons vu la longueur d'une courbe paramétrée, nous avons pris comme « élément de longueur » la norme du vecteur tangent. Il est donc naturel de prendre comme « élément de surface » une petite surface que l'on peut construire à partir des deux vecteurs tangents à la surface.

Au point  $\varphi(u_0, v_0)$ , nous avons les deux vecteurs tangents

$$T_u = \frac{\partial \varphi}{\partial u}(u_0, v_0) \quad T_v = \frac{\partial \varphi}{\partial v}(u_0, v_0). \quad (20.210)$$

L'élément de surface que nous pouvons construire à partir de ces deux vecteurs est la surface du parallélogramme, donnée par la norme du produit vectoriel :

$$dS = \|T_u \times T_v\|. \quad (20.211) \quad \text{EQooNYWSooZuyvPe}$$

L'aire de la surface donnée par  $\varphi : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  sera donc donnée par

$$\text{Aire}(\varphi(D)) = \iint_D \|T_u \times T_v\| du dv. \quad (20.212)$$

**Exemple 20.61.**

Calculons l'aire de la sphère. Les vecteurs tangents ont déjà été calculés aux équations (20.204) et (20.205) :

$$T_\theta = \begin{pmatrix} R \cos \theta \cos \varphi \\ R \cos \theta \sin \varphi \\ -R \sin \theta \end{pmatrix}, \quad T_\varphi = \begin{pmatrix} -R \sin \theta \sin \varphi \\ R \sin \theta \cos \varphi \\ 0 \end{pmatrix}. \quad (20.213)$$

Le produit vectoriel vaut

$$\begin{aligned} T_\theta \times T_\varphi &= \begin{vmatrix} e_x & e_y & e_z \\ R \cos \theta \cos \varphi & R \cos \theta \sin \varphi & -R \sin \theta \\ -R \sin \theta \sin \varphi & R \sin \theta \cos \varphi & 0 \end{vmatrix} \\ &= (R^2 \sin^2 \theta \cos \varphi) e_x + (R^2 \sin^2 \theta \sin \varphi) e_y \\ &\quad + (R^2 \cos \theta \sin \theta \cos^2 \varphi + R^2 \sin \theta \cos \theta \sin^2 \varphi) e_z. \end{aligned} \quad (20.214)$$

La norme demande quelques calculs et mises en évidences. Le résultat est :

$$\|T_\theta \times T_\varphi\| = R^2 \sin \theta. \quad \text{EqProdVectTTSPh} \quad (20.215)$$

L'aire de la sphère est donc donnée par

$$\text{Aire} = \int_0^{2\pi} d\varphi \int_0^\pi R^2 \sin \theta d\theta = 2\pi R^2 [-\cos \theta]_0^\pi = 4\pi R^2. \quad (20.216)$$

Il est bon de se souvenir que, en coordonnées sphériques,

$$\|T_\theta \times T_\varphi\| = R^2 \sin \theta. \quad (20.217)$$

Or nous savons que ce vecteur est dirigé dans le sens de  $e_r$  parce que ce dernier est le vecteur qui est constamment dirigé radialement. En coordonnées sphériques nous avons donc

$$T_\theta \times T_\varphi = R^2 \sin(\theta) e_r. \quad \text{EqNormalEnSph} \quad (20.218)$$

△

**Remarque 20.62.**

L'équation (20.215) donne l'élément de surface pour la sphère. Notez que cela est justement l'expression du jacobien des coordonnées sphériques. Cela n'est évidemment pas une coïncidence.

**Exemple 20.63.**

Nous pouvons donner l'aire du graphe d'une fonction quelconque. La surface est paramétrée par

$$\varphi(x, y) = \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix}. \quad (20.219)$$

Les vecteurs tangents sont

$$T_x = \begin{pmatrix} 1 \\ 0 \\ \partial_x f \end{pmatrix}, \quad T_y = \begin{pmatrix} 0 \\ 1 \\ \partial_y f \end{pmatrix}. \quad (20.220)$$

Le produit vectoriel est donné par

$$T_x \times T_y = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & \partial_x f \\ 0 & 1 & \partial_y f \end{vmatrix} = (-\partial_x f) e_x - (\partial_y f) e_y + e_z. \quad (20.221)$$

L'élément de surface est par conséquent

$$dS = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}, \quad (20.222)$$

et la surface du graphe sera

$$Aire = \iint_D \sqrt{\left(\frac{\partial f}{\partial x}(x, y)\right)^2 + \left(\frac{\partial f}{\partial y}(x, y)\right)^2 + 1} dx dy \quad (20.223)$$

△

### 20.9.2 Intégrale d'une fonction sur une surface

Si  $S$  est une surface dans  $\mathbb{R}^3$  paramétrée par

$$\begin{aligned} \varphi: D &\rightarrow \mathbb{R}^3 \\ (u, v) &\mapsto \varphi(u, v) \in S, \end{aligned} \quad (20.224)$$

et si  $f$  est une fonction  $f: \mathbb{R}^3 \rightarrow \mathbb{R}$  définie au moins sur  $S$ , l'intégrale de  $f$  sur  $S$  est logiquement définie par

$$\int_S f dS = \iint_D f(\varphi(u, v)) \|T_u(u, v) \times T_v(u, v)\| du dv \quad (20.225)$$

où  $T_u = \frac{\partial \varphi}{\partial u}$  et  $T_v = \frac{\partial \varphi}{\partial v}$ . La quantité

$$\|T_u(u, v) \times T_v(u, v)\| du dv \quad (20.226)$$

est appelé **élément de surface**.

Encore une fois, si on prend  $f = 1$ , alors on retrouve la surface de  $S$  :

$$\int_S dS = Aire(S). \quad (20.227)$$

#### Remarque 20.64.

Le nombre  $\int_S f dS$  ne dépend pas de le paramétrage choisie pour  $S$ .

### 20.9.3 Intégrale d'une 2-forme

Nous considérons  $\omega$ , une 2-forme différentielle sur  $\mathbb{R}^2$ .

#### Définition 20.65.

Si  $\omega_{(x,y)} = u(x, y) dx \wedge dy$  et si  $D$  est un ouvert de  $\mathbb{R}^2$  alors nous définissons

$$\int_D \omega = \int_D u(x, y) dx dy. \quad (20.228)$$

Nous voulons maintenant intégrer une 2-forme sur une surface dans  $\mathbb{R}^3$ . Soit  $S \subset \mathbb{R}^3$ , une surface orientée (c'est-à-dire que nous avons un choix continu d'un vecteur normal unitaire  $n$ ). Nous supposons de plus avoir un paramétrage  $\phi: D \rightarrow S$  de  $S$  avec  $D$  ouvert dans  $\mathbb{R}^2$  compatible avec l'orientation, c'est-à-dire que pour tout  $(t, s) \in D$ ,

$$n(\phi(t, s)) = \frac{\partial \phi}{\partial t}(t, s) \times \frac{\partial \phi}{\partial s}(t, s). \quad (20.229)$$

#### Définition 20.66.

Pour intégrer  $\omega$  sur  $S$  nous faisons

$$\int_S \omega = \int_D \phi^* \omega \quad (20.230)$$

où  $\phi^* \omega$  est de la forme  $F(t, s) dt \wedge ds$ .

Montrons ce que cela fait. Soient  $u, v$  des vecteurs de  $D$  et calculons

$$(\phi^*\omega)(u, v) = \omega(d\phi(u), d\phi(v)) \quad (20.231a)$$

$$= \omega\left(u_1 \frac{\partial \phi}{\partial x_1} + u_2 \frac{\partial \phi}{\partial x_2}, v_1 \frac{\partial \phi}{\partial x_1} + v_2 \frac{\partial \phi}{\partial x_2}\right). \quad (20.231b)$$

Les termes en  $u_1v_1$  et  $u_2v_2$  sont nuls ; par exemple :

$$\omega\left(u_1 \frac{\partial \phi}{\partial x_1}, v_1 \frac{\partial \phi}{\partial x_1}\right) = u_1v_1\omega\left(\frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_1}\right) = 0 \quad (20.232)$$

parce que  $\omega$  est antisymétrique. Il nous reste donc

$$(\phi^*\omega)(u, v) = (u_1v_2 - u_2v_1)\omega\left(\frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_2}\right) \quad (20.233a)$$

$$= \omega\left(\frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_2}\right)(dt \wedge ds)(u, v). \quad (20.233b)$$

Cette dernière ligne est bien de la forme  $\phi^*\omega = F(t, s)dt \wedge ds$ .

## 20.10 Flux d'un champ de vecteurs à travers une surface

Nous voulons construire un moulin à eau. Comment placer les pales pour maximiser le travail de la pression de l'eau ? On n'a pas attendu l'invention du calcul intégral pour répondre à cette question. Trois paramètres rentrent en ligne de compte :

- (1) plus il y a d'eau, plus ça pousse ;
- (2) plus la surface de la pale est grande, plus on va utiliser d'eau ;
- (3) plus la pale est perpendiculaire au courant, plus on va en profiter.

Nous voyons sur la figure 20.2 que lorsque la pale du moulin est inclinée, non seulement elle prend moins d'eau sur elle, mais qu'en plus elle la prend avec un moins bon angle : une partie de la force ne sert pas à la faire tourner.

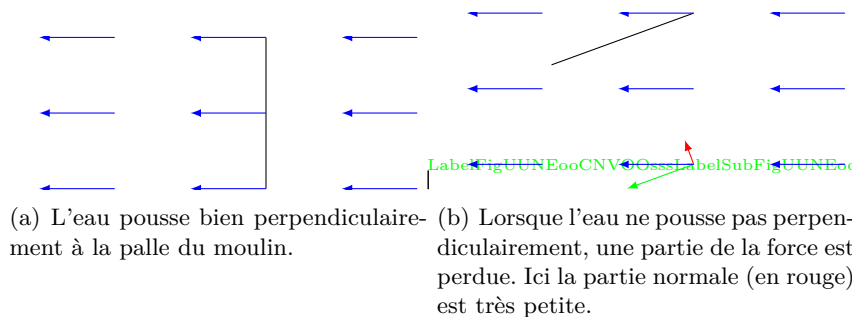


FIGURE 20.2: La partie rouge de la force est perdue si l'eau ne pousse pas perpendiculairement. De plus lorsque la pale est inclinée, elle prend moins d'eau sur elle.

L'idée du flux d'un champ de vecteurs à travers une surface est de savoir quelle est la quantité « utile » de vecteurs qui traverse la surface. Ce sera simplement l'intégrale sur la surface de la composante du champ de vecteurs normale à la surface. Il reste deux problèmes à régler : le premier est de savoir quel est le vecteur normal à la surface, et le second est de savoir comment « sélectionner » la composante normale d'un champ de vecteurs  $F$ .

Le problème de trouver un vecteur normal est résolu par le produit vectoriel des vecteurs tangents. Si la surface est donnée par  $\varphi: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ , les vecteurs tangents sont  $T_u = \partial_u \varphi(u, v)$  et  $T_v = \partial_v \varphi(u, v)$ . Le normal de norme 1 est donné par :

$$n(u, v) = \frac{T_u \times T_v}{\|T_u \times T_v\|}. \quad (20.234)$$

Si  $p$  est un point de la surface  $\varphi(D)$ , la composante de  $F(p)$  qui est normale à la surface au point  $p$  est donnée par le produit scalaire

$$F(p)_\perp = F(p) \cdot n(p). \quad (20.235)$$

C'est ce nombre là que nous intégrons sur la surface.

**Définition 20.67.**

Le **flux du champ de vecteurs** à travers la surface  $S = \varphi(D)$  est

$$\int F \cdot dS = \int F \cdot n \, dudv. \quad (20.236)$$

Une petite simplification se produit lorsqu'on veut calculer effectivement cette intégrale. En effet  $F \cdot n$  est, en soi, une fonction sur  $S$ . Pour l'intégrer, il faut donc la multiplier par  $\|T_u \times T_v\|$  (c'est la définition de l'intégrale d'une fonction sur une surface). Donc, étant donné que  $n = (T_u \times T_v)/\|T_u \times T_v\|$ , nous avons

$$\int F \cdot dS = \iint_D F(\varphi(u, v)) \cdot (T_u \times T_v) \, dudv \quad (20.237)$$

où  $T_u = \frac{\partial \varphi}{\partial u}$  et  $T_v = \frac{\partial \varphi}{\partial v}$ .

**Exemple 20.68.**

Soit le champ de vecteurs

$$F = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}. \quad (20.238)$$

Calculons son flux au travers de la sphère de rayon  $R$ .

Nous choisissons de paramétrer la sphère en coordonnées sphériques avec  $\phi(\theta, \varphi)$ . Nous pouvons reprendre le résultat (20.218) :

$$T_\theta \times T_\varphi = R^2 \sin(\theta). \quad (20.239)$$

Nous savons aussi que

$$F(\phi(\theta, \varphi)) = R e_r. \quad (20.240)$$

L'intégrale à calculer est donc

$$I = \int_0^\pi d\theta \int_0^{2\pi} d\varphi \, 2e_r \cdot (R^2 \sin(\theta) e_r). \quad (20.241)$$

Vu que le produit scalaire  $e_r \cdot e_r$  vaut 1, nous calculons

$$I = 4\pi R^2 \int_0^\pi \sin(\theta) d\theta = 8\pi R^2. \quad (20.242)$$

△

**Exemple 20.69.**

Calculons le flux du champ de force de gravitation d'une masse au travers de la sphère de centre  $R$  centrée autour la masse. À un coefficient constant près, le champ vaut

$$G(r, \theta, \varphi) = \frac{1}{r^2} e_r. \quad (20.243)$$

Sur la sphère de rayon  $R$ , nous avons

$$G(\phi(\theta, \varphi)) = \frac{1}{R^2} e_r. \quad (20.244)$$

L'intégrale est donc

$$\int_0^\pi d\theta \int_0^{2\pi} \frac{1}{R^2} e_r \cdot (R^2 \sin(\theta) e_r) d\varphi = 8\pi. \quad (20.245)$$

Ce flux ne dépend pas de  $R$ .

△

**Exemple 20.70.**

Soit  $S$  le disque de rayon 5 placé horizontalement à la hauteur 12. Calculer le flux du champ de vecteurs

$$F(x, y, z) = xe_x + ye_y + ze_z. \quad (20.246)$$

Les équations de la surface sont  $z = 12$ ,  $x^2 + y^2 \leq 25$ . Nous prenons le paramétrage en coordonnées cylindriques :

$$\varphi(r, \theta) = \begin{pmatrix} r \cos(\theta) \\ r \sin(\theta) \\ 12 \end{pmatrix}. \quad (20.247)$$

Les vecteurs tangents sont

$$T_r = \frac{\partial \varphi}{\partial r} = \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix} \quad T_\theta = \frac{\partial \varphi}{\partial \theta} = \begin{pmatrix} -r \sin \theta \\ r \cos \theta \\ 0 \end{pmatrix}. \quad (20.248)$$

Le vecteur normal est alors

$$T_r \times T_\theta = re_z. \quad (20.249)$$

Sur la surface, le champ de vecteurs s'écrit

$$F(\varphi(r, \theta)) = r \cos(\theta)e_x + r \sin(\theta)e_y + 12e_z. \quad (20.250)$$

Par conséquent

$$F \cdot (T_r \times T_\theta) = 12r. \quad (20.251)$$

L'intégrale à calculer est

$$\begin{aligned} \int_0^5 dr \int_0^{2\pi} 12r d\theta &= 12 \cdot 2\pi \int_0^5 r dr \\ &= \frac{25}{2} 24\pi \\ &= 300\pi. \end{aligned} \quad (20.252)$$

△

## 20.11 Divergence, Green, Stokes

Le théorème de Stokes (et ses variations) peut se voir comme une généralisation du théorème fondamental du calcul différentiel et intégral qui stipule que

$$\int_a^b f'(x) dx = f(b) - f(a)$$

c'est-à-dire qui relie l'intégrale de  $f'$  sur  $I = [a, b]$  aux valeurs de  $f$  sur le bord  $\partial I = \{a, b\}$ . Le signe  $-$  qui apparait vient de l'orientation ; celle-ci requiert de la prudence dans l'utilisation des théorèmes.

Voici, pour votre culture générale, un énoncé général :

ThoATsPuzF

**Théorème 20.71.**

Si  $M$  est une variété orientable de dimension  $n$  avec un bord noté  $\partial M$ , alors pour toute forme différentielle  $\omega$  de degré  $n - 1$  on a

$$\int_M d\omega = \int_{\partial M} \omega.$$

où  $d\omega$  désigne la différentielle extérieure de  $\omega$ .

Nous allons maintenant voir quelques cas particuliers.

Une des nombreuses formes du théorème de Stokes (théorème 20.71) est que si la forme différentielle  $\omega$  est exacte alors son intégrale est facile.

**Théorème 20.72** ([1]).

Si  $\gamma$  est une chemin de classe  $C^1$  dans un ouvert  $\Omega$  et si  $\omega$  est la forme différentielle exacte  $\omega = df$ , alors

$$\int_{\gamma} df = f(\gamma(1)) - f(\gamma(0)). \quad (20.253)$$

*Démonstration.* C'est une application du lemme 12.306 et du théorème fondamental du calcul intégral 14.288.  $\square$

**20.11.1 Théorème de la divergence**

Si nous considérons une surface dans  $\mathbb{R}^n$  et un champ de vecteurs, il est bon de se demander quelle « quantité de vecteurs » traverse la surface. Soit  $D$ , un ouvert borné de  $\mathbb{R}^n$  telle que  $\partial D$  soit une variété de dimension  $n-1$ , et  $G$ , un champ de vecteurs défini sur  $\bar{D}$ . Afin de compter combien de  $G$  traverse  $\partial D$ , il faudra faire en sorte de ne considérer que la composante de  $G$  normale à  $\partial D$  : pas question d'intégrer par exemple la norme de  $G$  sur  $\partial D$ .

Comme nous le savons, la composante du vecteur  $v$  dans la direction  $w$  est le produit scalaire  $v \cdot 1_w$  où  $1_w$  est le vecteur de norme 1 dans la direction  $w$ . Nous allons donc introduire le concept de vecteur normal extérieur.

**Définition 20.73.**

Soit  $x \in \partial D$  et  $\nu \in \mathbb{R}^n$ , nous disons que  $\nu$  est un **vecteur normal extérieur** de  $\partial D$  si

- (1)  $\langle \nu, v \rangle = 0$  pour tout vecteur tangent  $v$  à  $\partial D$  au point  $x$ . Pour rappel,  $\partial D$  étant une variété de dimension  $n-1$ , il y a  $n-1$  tels vecteurs  $v$  linéairement indépendants.
- (2) Il existe un  $\delta > 0$  tel que  $\forall t \in ]0, \delta[$ , nous avons  $x + t\nu \notin \bar{D}$  et  $x - t\nu \in D$ .

Nous pouvons maintenant définir le concept de flux. Soit  $D \subset \mathbb{R}^n$  tel que  $\partial D$  soit une variété de dimension  $n-1$  qui admette un vecteur normal extérieur  $\nu(x)$  en chaque point. Soit aussi  $G: \bar{D} \rightarrow \mathbb{R}^n$ , un champ de vecteur de classe  $C^1$ . Le **flux** de  $G$  au travers de  $\partial D$  est le nombre

$$\int_{\partial D} \langle G(x), \nu(x) \rangle d\sigma(x). \quad (20.254)$$

Cette intégrale est en général très compliquée à calculer parce qu'il faut trouver le champ de vecteur normal, puis un paramétrage de la surface  $\partial D$  et ensuite appliquer la méthode décrite au point 20.8.

Heureusement, il y a un théorème qui nous permet de calculer plus facilement : sans devoir trouver de vecteurs normaux.

Il n'est pas plus contraignant d'énoncer ce théorème dans le cadre d'une hypersurface de  $\mathbb{R}^n$ , ce que nous faisons donc :

**Théorème 20.74** (Formule de la divergence).

Soit  $D$  un ouvert borné de  $\mathbb{R}^n$  dont le bord est « assez régulier par morceaux », c'est-à-dire :

$$\partial D = A_1 \cup \dots \cup A_p \cup N \quad (20.255)$$

où

- (1)  $A_1, \dots, A_p, N$  sont deux à deux disjoints,
- (2) pour tout  $i \leq p$ ,  $A_i$  est un ouvert relatif d'une certaine variété  $M_i$  de dimension  $(n-1)$
- (3)  $\bar{A}_i \subset M_i$
- (4)  $N$  est un compact contenu dans une réunion finie de variétés de dimensions  $(n-2)$ .

Supposons également qu'en chaque point de  $A_1 \cup \dots \cup A_p$  il existe un vecteur normal extérieur  $\nu$ .

Si  $G$  est un champ de vecteurs de classe  $C^1$  sur  $\bar{D}$  alors

$$\int_D \nabla \cdot G = \sum_{i=1}^p \int_{A_i} \langle G, \nu \rangle. \quad (20.256)$$

L'intégrale du membre de gauche est l'intégrale sur un ouvert d'une simple fonction.



### 20.11.2 Lacets et homotopie

DEFooQZMSooYYkGDv

**Définition 20.75** (chemin, lacet[234, 654]).

Plusieurs notions autour de chemins dans un espace topologique  $X$ .

- (1) Un **chemin** dans un espace vectoriel  $X$  est une application continue  $\gamma: [a, b] \rightarrow X$  avec  $a < b$ .
- (2) Un **lacet** dans  $X$  est un chemin  $\gamma: [a, b] \rightarrow X$  tel que  $\gamma(a) = \gamma(b)$ .
- (3) Un chemin  $\gamma: [0, 1] \rightarrow \mathbb{R}^n$  est **régulier** si il est  $C^1$  et si  $\gamma'(t) \neq 0$  pour tout  $t$ .
- (4) Le chemin  $\gamma: [a, b] \rightarrow X$  est un **chemin de Jordan** si il est injectif.
- (5) Un lacet  $\gamma: [a, b] \rightarrow X$  est un **lacet de Jordan** si  $\gamma: [a, b[ \rightarrow X$  est injective.
- (6) Une **courbe de Jordan** est l'image d'un lacet de Jordan.
- (7) Une **courbe simple** est l'image d'un chemin de Jordan.

L'homotopie sera définie en 20.77.

#### 20.76.

Un chemin de Jordan peut évidemment être vu comme une application  $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$  telle que  $\gamma(0) = \gamma(2\pi)$ . En particulier il n'est jamais mauvais de se rappeler qu'on peut choisir un paramétrage normal par la proposition 21.46.

DEFooHJQTooYUFcee

**Définition 20.77** (homotopie de lacets).

Soit un espace topologique  $X$ . Les lacets<sup>17</sup>  $\gamma_0, \gamma_1: [a, b] \rightarrow X$  sont **homotopes** dans  $X$  si il existe une application continue

$$H: [0, 1] \times [a, b] \rightarrow X \quad (20.257)$$

telle que

- (1) Pour tout  $s \in [0, 1]$ , l'application

$$\begin{aligned} \Gamma_s: [a, b] &\rightarrow X \\ t &\mapsto H(s, t) \end{aligned} \quad (20.258)$$

est un lacet dans  $X$ .

- (2) Nous avons  $\Gamma_0 = \gamma_0$  et  $\Gamma_1 = \gamma_1$ .

L'application  $H$  est l'homotopie entre  $\gamma_0$  et  $\gamma_1$ .

DEFooLXD'TooDPgxqL

**Définition 20.78** (Homotopie à extrémités fixées[234]).

Soit un espace topologique  $X$ . Soient des chemins  $\gamma_0, \gamma_1: [a, b] \rightarrow X$ . Nous supposons que leurs extrémités soient égales, et nous les notons  $p, q$  :  $\gamma_0(a) = \gamma_1(a) = p$  et  $\gamma_0(b) = \gamma_1(b) = q$ . Nous disons que  $\gamma_0$  et  $\gamma_1$  sont **homotopes avec les extrémités fixées** si il existe une application continue  $H: [0, 1] \times [a, b] \rightarrow X$  telle que

- (1) les extrémités sont fixées :  $H(s, a) = p$  et  $H(s, b) = q$  pour tout  $s \in [0, 1]$ .
- (2) pour tout  $t \in [a, b]$ , nous avons  $H(0, t) = \gamma_0(t)$  et  $H(1, t) = \gamma_1(t)$ .

### 20.11.3 Formule de Green

La formule de Green est un cas particulier du théorème de la divergence dans le cas  $n = 2$ , légèrement reformulé.

THOooQSWMooAZasTl

**Théorème 20.79.**

Soit  $D \subset \mathbb{R}^2$  ouvert borné tel que son bord est la réunion finie d'un certain nombre de chemins

---

17. Définition 20.75.

de classe  $C^1$  de Jordan réguliers. Supposons qu'en chaque point de son bord,  $D$  possède un vecteur normal unitaire extérieur  $\nu$ . Soient  $P$  et  $Q$  deux fonctions réelles de classe  $C^1$  sur  $\bar{D}$ . Alors

$$\iint_D (\partial_x Q - \partial_y P) dx dy = \oint_{\partial D} P dx + Q dy \quad (20.259) \quad \text{EqYLbISaV}$$

où chaque chemin  $\gamma$  formant le bord de  $D$  est orienté de sorte que  $T\nu = \frac{\dot{\gamma}}{\|\dot{\gamma}\|}$  où  $T$  représente la rotation d'angle  $+\frac{\pi}{2}$ .

Justifions le fait que cela soit un cas particulier de la formule de Stokes du théorème 20.71. Nous considérons la forme différentielle

$$\omega = P dx + Q dy, \quad (20.260)$$

et sa différentielle

$$d\omega = \sum_i d\omega_i \wedge dx_i \quad (20.261a)$$

$$= \left( \frac{\partial P}{\partial x} dx + \frac{\partial P}{\partial y} dy \right) \wedge dx + \left( \frac{\partial Q}{\partial x} dx + \frac{\partial Q}{\partial y} dy \right) \wedge dy \quad (20.261b)$$

$$= \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx \wedge dy. \quad (20.261c)$$

Intégrons cette forme  $d\omega$  sur le domaine ouvert  $D$  que nous paramétrons de façon triviale par

$$\begin{aligned} \varphi: D &\rightarrow \mathbb{R}^2 \\ (u, v) &\mapsto (u, v). \end{aligned} \quad (20.262)$$

Ce que nous avons est

$$\iint_D d\omega = \iint_D d\omega_{(u,v)} \left( \frac{\partial \varphi}{\partial u}, \frac{\partial \varphi}{\partial v} \right) du dv \quad (20.263) \quad \text{EqKYiFEGF}$$

Nous avons aussi  $T_u = \frac{\partial \varphi}{\partial u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  et  $T_v = \frac{\partial \varphi}{\partial v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  et donc

$$(dx \wedge dy)(T_u, T_v) = dx(T_u)dy(T_v) - dx(T_v)dy(T_u) = 1 - 0 = 1. \quad (20.264)$$

L'intégrale (20.263) se développe donc en

$$\iint_D d\omega = \iint_D \left( \frac{\partial Q}{\partial x}(u, v) - \frac{\partial P}{\partial y}(u, v) \right) (dx \wedge dy)(T_u, T_v) du dv = \iint_D \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) du dv. \quad (20.265)$$

Par conséquent la formule de Stokes nous donne la formule (20.259).

La formule de Green nous permet de calculer l'aire de la surface délimitée par une courbe fermée en termes de l'intégrale d'une forme bien choisie le long du contour. Pour cela nous prenons la forme

$$\omega = -\frac{y}{2} dx + \frac{x}{2} dy, \quad (20.266) \quad \text{EqZNXYYMQb}$$

de telle sorte que  $\partial_x Q - \partial_y Q = 1$  et que

$$\iint_D d\omega = \iint_D du dv = S, \quad (20.267)$$

et au final l'aire est donnée par

$$S = \int_{\partial D} \left( -\frac{y}{2} dx + \frac{x}{2} dy \right). \quad (20.268)$$

Lorsque le bord de  $D$  est paramétré par

$$\begin{aligned}\gamma: [a, b] &\rightarrow \mathbb{R}^2 \\ u &\mapsto \begin{pmatrix} x(u) \\ y(u) \end{pmatrix},\end{aligned}\tag{20.269}$$

nous avons

$$(Pdx + Qdy)\gamma'(u) = Px' + Qy',\tag{20.270}$$

et alors

$$\int_{\partial D} Pdx + Qdy = \int_a^b P(x(u), y(u))x'(u) + Q(x(u), y(u))y'(u)du.\tag{20.271}$$

En ce qui concerne l'aire de la surface, nous prenons les  $P$  et  $Q$  de la forme 20.266 :

$$S = \frac{1}{2} \int_a^b \left( -y(u)x'(u) + x(u)y'(u) \right) du.\tag{20.272}$$

#### 20.11.4 Formule de Stokes

La formule de Stokes est la version classique, qui permet d'exprimer la circulation d'un champ de vecteur le long d'une courbe de  $\mathbb{R}^3$  comme le flux de son rotationnel à travers n'importe quelle surface dont le bord est la courbe. La version présentée ici suppose que la surface peut se paramétrer en un seul morceau :

**Théorème 20.80** (Formule de Stokes).

Soit  $F : W \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$  un paramétrage (carte) d'une surface dans  $\mathbb{R}^3$ , supposée de classe  $C^2$ . Soit  $D$  un ouvert de  $\mathbb{R}^2$  vérifiant les hypothèses de la formule de Green, et tel que  $\bar{D} \subset W$ . Soit  $G$  un champ de vecteurs de classe  $C^1$  défini sur  $F(\bar{D})$ , et soit  $N$  le champ normal unitaire donné par le paramétrage

$$N = \frac{\frac{\partial F}{\partial u} \wedge \frac{\partial F}{\partial v}}{\left\| \frac{\partial F}{\partial u} \wedge \frac{\partial F}{\partial v} \right\|}\tag{20.273}$$

alors

$$\iint_{F(D)} \langle \nabla \times G, N \rangle d\sigma_F = \int_{F(\partial D)} G\tag{20.274}$$

où les chemins formant le bord  $\partial D$  sont orientés comme dans le théorème de Green.

Notons, juste pour avoir une bonne nouvelle de temps en temps, que

$$d\sigma_F = \left\| \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\| dudv,\tag{20.275}$$

mais cette norme apparaît exactement au dénominateur de  $N$ . Il ne faut donc pas la calculer parce qu'elle se simplifie.

Sous forme un peu plus physicienne<sup>18</sup>, la formule (20.274) s'écrit

$$\int_{F(D)} \langle \nabla \times G, N(x) \rangle d\sigma_F(x) = \int_{F(\gamma)} \langle G, T \rangle ds\tag{20.276}$$

où  $T$  est le vecteur unitaire tangent à  $F(\gamma)$ .

---

18. et surtout plus explicite.

### 20.11.4.1 Quelle est la bonne orientation ?

Le signe du vecteur normal  $N$  dépend du choix de l'ordre des coordonnées dans la carte. Supposons que je veuille paramétrer la surface  $x^2 + y^2 = 1, z = 1$ . Nous prenons naturellement comme carte le cercle  $C$  de rayon 1 dans  $\mathbb{R}^2$  et la carte

$$F(r, \theta) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ 1 \end{pmatrix}. \quad (20.277)$$

Mais nous aurions aussi pu mettre les coordonnées  $r$  et  $\theta$  dans l'autre ordre :

$$\tilde{F}(\theta, r) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ 1 \end{pmatrix}. \quad (20.278)$$

Les vecteurs normaux ne sont pas les mêmes : la carte  $F$  donnera  $\partial_r F \times \partial_\theta F$ , tandis que l'autre donnera  $\partial_\theta \tilde{F} \times \partial_r \tilde{F}$ . Le signe change !

Il faut savoir laquelle choisir. Le cercle  $C \subset \mathbb{R}^2$  a une orientation donnée par le théorème de Green. Nous choisissons l'ordre des coordonnées pour que  $1_\theta$  et  $1_r$  soient dans la même orientation que les vecteurs  $\nu$  et  $T$  tels que donnés par le théorème de Green, et tels que dessinés sur la figure 20.3.

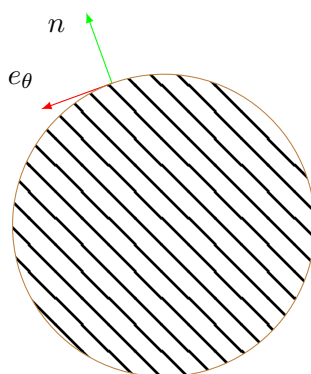


FIGURE 20.3: L'orientation sur le cercle. Si nous les prenons dans l'ordre, les vecteurs  $(1_r, 1_\theta)$  ont la même orientation que celle donnée par les vecteurs  $(\nu, T)$  donnés par la convention de Green.

Plus généralement, nous choisissons l'ordre des coordonnées  $u$  et  $v$  pour que la base  $(1_u, 1_v)$  ait la même orientation que  $(\nu, T)$  où  $T$  a le sens convenu dans le théorème de Green.

## 20.12 Résumé des intégrales vues

Nous sommes maintenant capables de revoir tous les types d'intégrales vues jusqu'ici de façon très cohérentes. Nous commencerons par les intégrales de fonctions et nous ferons ensuite les intégrales de champs de vecteurs.

### 20.12.1 L'intégrale d'une fonction sur les réels

Si  $f: [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$  est une fonction usuelle, son intégrale est

$$\int_a^b f(x) dx = F(b) - F(a) \quad (20.279)$$

où  $F$  est une primitive de  $f$ .

### 20.12.2 Intégrale d'une fonction sur un chemin

Si  $f$  est une fonction sur  $\mathbb{R}^3$  et si  $\sigma: [a, b] \rightarrow \mathbb{R}^3$  est un chemin dans  $\mathbb{R}^3$ , l'intégrale de  $f$  sur  $\sigma$  est, par définition,

$$\int f d\sigma = \int_a^b f(\sigma(t)) \|\sigma'(t)\| dt. \quad (20.280)$$

### 20.12.3 Intégrale d'une fonction sur une surface

Nous devons paramétrer la surface  $S$  par une application  $\varphi: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ . À partir d'un tel paramétrage, nous construisons un élément de surface en prenant le produit vectoriel des deux vecteurs tangents :

$$dS = \frac{\partial \varphi}{\partial u} \times \frac{\partial \varphi}{\partial v} du dv. \quad (20.281)$$

L'intégrale est

$$\int f dS = \iint_D f(\varphi(u, v)) \left\| \frac{\partial \varphi}{\partial u} \times \frac{\partial \varphi}{\partial v} \right\| du dv. \quad (20.282)$$

Il ne faut pas rajouter de jacobien : la norme du produit vectoriel *est* le jacobien.

#### Remarque 20.81.

La formule (20.282) est autant valable pour des surfaces dans  $\mathbb{R}^2$  que dans  $\mathbb{R}^3$ . Si nous considérons une surface dans  $\mathbb{R}^2$ , nous la voyons dans  $\mathbb{R}^3$  en ajoutant un zéro comme troisième composante.

#### Exemple 20.82.

Les coordonnées polaires sont données par

$$\varphi(r, \theta) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \\ 0 \end{pmatrix}. \quad (20.283)$$

Les vecteurs tangents à ce paramétrage sont

$$T_r = \frac{\partial \varphi}{\partial r} = \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix}, \quad T_\theta = \frac{\partial \varphi}{\partial \theta} = \begin{pmatrix} -r \sin \theta \\ r \cos \theta \\ 0 \end{pmatrix}. \quad (20.284)$$

Le vecteur normal est

$$\frac{\partial \varphi}{\partial r} \times \frac{\partial \varphi}{\partial \theta} = \begin{vmatrix} e_x & e_y & e_z \\ \cos \theta & \sin \theta & 0 \\ -r \sin \theta & r \cos \theta & 0 \end{vmatrix} = r e_z. \quad (20.285)$$

Nous trouvons donc que l'élément de surface est la norme de  $r e_z$ , c'est-à-dire  $r$ , le jacobien connu.  $\triangle$

### 20.12.4 Intégrale d'une fonction sur un volume

Si  $V$  est un volume dans  $\mathbb{R}^3$ , nous effectuons la même procédure : nous trouvons un paramétrage, et nous formons un élément de volume avec les vecteurs tangents de le paramétrage. Nous avons donc un volume déterminé par l'application

$$\varphi: D \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (20.286)$$

et ses trois vecteurs tangents

$$\begin{aligned} T_u &= \frac{\partial \varphi}{\partial u} \\ T_v &= \frac{\partial \varphi}{\partial v} \\ T_w &= \frac{\partial \varphi}{\partial w}. \end{aligned} \quad (20.287)$$

Comment former un volume avec trois vecteurs ? Réponse : le produit mixte. L'intégrale de  $f$  sur  $V$  sera

$$\int f dV = \iiint_D f(\varphi(u, v, w)) \left\| \frac{\partial \varphi}{\partial u} \cdot \left( \frac{\partial \varphi}{\partial v} \times \frac{\partial \varphi}{\partial w} \right) \right\| du dv dw. \quad (20.288)$$

Encore une fois, le produit mixte *est* le jacobien. Prenons les coordonnées sphériques :

$$\begin{aligned} x(r, \theta, \varphi) &= r \sin(\theta) \cos(\varphi) \\ y(r, \theta, \varphi) &= r \sin(\theta) \sin(\varphi) \\ z(r, \theta, \varphi) &= r \cos(\theta) \end{aligned} \quad (20.289)$$

Les trois vecteurs tangents seront

$$T_r = \begin{pmatrix} \frac{\partial x}{\partial r} \\ \frac{\partial y}{\partial r} \\ \frac{\partial z}{\partial r} \end{pmatrix} = \begin{pmatrix} \sin(\theta) \cos(\varphi) \\ \sin(\theta) \sin(\varphi) \\ \cos(\theta) \end{pmatrix} \quad (20.290a)$$

$$T_\theta = \begin{pmatrix} \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \theta} \end{pmatrix} = \begin{pmatrix} r \cos(\theta) \cos(\varphi) \\ r \cos(\theta) \sin(\varphi) \\ -r \sin(\theta) \end{pmatrix} \quad (20.290b)$$

$$T_\varphi = \begin{pmatrix} \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} -r \sin(\theta) \sin(\varphi) \\ r \sin(\theta) \cos(\varphi) \\ 0 \end{pmatrix} \quad (20.290c)$$

Nous avons vu que le produit mixte revient à mettre toutes les composantes dans une matrice. Ici nous avons donc

$$\frac{\partial \phi}{\partial r} \cdot \left( \frac{\partial \phi}{\partial \theta} \times \frac{\partial \phi}{\partial \varphi} \right) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} & \frac{\partial z}{\partial r} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} & \frac{\partial z}{\partial \theta} \\ \frac{\partial x}{\partial \varphi} & \frac{\partial y}{\partial \varphi} & \frac{\partial z}{\partial \varphi} \end{vmatrix} \quad (20.291)$$

Cela est précisément le jacobien dont nous parlions plus haut.

### 20.12.5 Conclusion pour les fonctions

Lorsque nous intégrons une fonction sur un chemin, une surface ou un volume, la technique est toujours la même :

- (1) Trouver un paramétrage à une, deux ou trois variables.
- (2) Dériver le paramétrage par rapport à ses variables.
- (3) Construire un élément de longueur, surface ou volume à partir des vecteurs que l'on a. Cela se fait en prenant la norme, le produit vectoriel ou le produit mixte.

### 20.12.6 Circulation d'un champ de vecteurs

Pour les champs de vecteurs, nous faisons la même chose, mais au lieu de *multiplier* par l'élément de longueur ou de surface, nous prenons le produit scalaire. Si nous considérons la courbe paramétrée  $\sigma: [a, b] \rightarrow \mathbb{R}^3$  et le champ de vecteurs  $F$ , nous avons donc

$$\int_\sigma F = \int F \cdot d\sigma = \int_a^b F(\sigma(t)) \cdot \sigma'(t) dt. \quad (20.292)$$

### 20.12.7 Flux d'un champ de vecteurs

Si la surface  $S \subset \mathbb{R}^3$  est paramétrée par

$$\begin{aligned} \phi: D \subset \mathbb{R}^2 &\rightarrow \mathbb{R}^3 \\ (u, v) &\mapsto \phi(u, v), \end{aligned} \quad (20.293)$$

et si  $F$  est un champ de vecteurs, alors on a

$$\int_S F = \int_S F \cdot dS = \iint_D F(\phi(u, v)) \cdot \left( \frac{\partial \phi}{\partial u} \times \frac{\partial \phi}{\partial v} \right) dudv. \quad \text{EqResIntFluxPhi} \quad (20.294)$$

### 20.12.8 Conclusion pour les champs de vecteurs

La circulation et le flux ne représentent pas tout à fait la même chose. En effet pour la circulation, nous sélectionnons la composante *tangente* à la courbe, c'est-à-dire la partie du vecteurs qui « circule » le long de la courbe. Une force perpendiculaire au mouvement ne travaille pas.

La situation est exactement le contraire pour le flux. Étant donné que le vecteur

$$\frac{\partial \phi}{\partial u} \times \frac{\partial \phi}{\partial v} \quad (20.295)$$

est normal à la surface, le fait de prendre le produit scalaire du champ de vecteurs avec lui sélectionne la composante *normale* à la surface, c'est-à-dire la partie du vecteur qui traverse la surface.

### 20.12.9 Attention pour les surfaces fermées !

Si nous considérons une surface fermée, il faut faire attention à choisir une *orientation*. Les vecteurs normaux doivent soit tous pointer vers l'intérieur soit tous vers l'extérieur. En effet, en tant que vecteur normal, nous avons choisi de prendre

$$T_u \times T_v. \quad (20.296)$$

Mais le vecteur  $T_v \times T_u$  est tout aussi normal ! Il n'y a pas à priori de façon standard pour choisir l'un ou l'autre. Il faut juste être cohérent : il faut que si on divise la surface en plusieurs morceaux, tous les vecteurs pointent dans le même sens.

Notez que si vous faites un choix et que votre voisin fait le choix inverse, vous obtiendrez des réponses qui diffèrent d'un signe. Sans plus de précisions<sup>19</sup>, les deux réponses sont correctes.

Un exemple de ce problème est donné dans l'exemple 20.83.

#### Exemple 20.83.

Calculer le flux du champ de vecteurs

$$F(x, y, z) = e_x \quad (20.297)$$

au travers du cylindre de rayon  $R$  et de hauteur  $h$  autour de l'axe  $z$ .

Même question si le cylindre est autour de l'axe  $x$ .

Remarque : ces cylindres sont considérés *avec* leur « couvercles ».

Un paramétrage du cylindre autour de l'axe  $z$  est

$$\phi(\theta, z) = \begin{pmatrix} R \cos \theta \\ R \sin \theta \\ z \end{pmatrix}. \quad (20.298)$$

Les vecteurs tangents sont

$$T_\theta = \begin{pmatrix} -R \sin \theta \\ R \cos \theta \\ 0 \end{pmatrix}, T_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (20.299)$$

Le vecteur normal est donc

$$T_\theta \times T_z = R \cos(\theta) e_x + R \sin(\theta) e_y. \quad (20.300)$$

19. Il faudrait définir ce qu'est une surface *orientable* et choisir une orientation.

C'est un vecteur dirigé vers l'extérieur.

Le champ de vecteurs considéré est constant :  $F(\theta, z) = e_x$ . Nous avons donc

$$F(\theta, z) \cdot (T_\theta \times T_z) = R \cos(\theta) \quad (20.301)$$

et le flux vaut

$$\Phi = \int_0^{2\pi} d\theta \int_0^h R \cos(\theta) dz = 0. \quad (20.302)$$

En ce qui concerne les couvercles haut au bas, ils sont paramétrés par

$$\phi_1(r, \theta) = \begin{pmatrix} R \cos(\theta) \\ R \sin(\theta) \\ h \end{pmatrix}, \phi_2(r, \theta) = \begin{pmatrix} R \cos(\theta) \\ R \sin(\theta) \\ 0 \end{pmatrix}. \quad (20.303)$$

Les vecteurs normaux correspondants sont dans la direction de  $e_z$ , de façon que le produit scalaire avec  $F(r, \theta)$  soit nul. Le flux total est donc nul.

Regardons maintenant le cylindre le long de l'axe  $x$ . Un paramétrage est

$$\phi(\theta, x) = \begin{pmatrix} x \\ R \cos(\theta) \\ R \sin(\theta) \end{pmatrix}, \quad (20.304)$$

et les vecteurs tangents sont

$$T_\theta = \begin{pmatrix} 0 \\ -R \sin \theta \\ R \cos \theta \end{pmatrix}, T_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \quad (20.305)$$

Le vecteur normal est alors donné par

$$T_\theta \times T_x = R \cos(\theta) e_y + R \sin(\theta) e_z. \quad (20.306)$$

Nous avons par conséquent  $F(\theta, x) \cdot (T_\theta \times T_x) = 0$ . Pas de flux par le côté du cylindre.

Regardons les « couvercles ». Le premier est donné par le paramétrage

$$\phi_1(r, \theta) = \begin{pmatrix} 0 \\ r \cos(\theta) \\ r \sin(\theta) \end{pmatrix}. \quad (20.307)$$

Le vecteur normal serait  $T_r \times T_\theta = r e_x$ , et le flux

$$\Phi = \int_0^{2\pi} d\theta \int_0^R r dr = \pi R^2. \quad (20.308)$$

Le second couvercle est donné par

$$\phi_2(r, \theta) = \begin{pmatrix} h \\ r \cos(\theta) \\ r \sin(\theta) \end{pmatrix}. \quad (20.309)$$

Le vecteur normal est encore  $r e_x$ , et le flux est à nouveau  $\pi R^2$ .

Le flux total serait donc  $2\pi R^2$ .

Cela n'est pas possible parce que tous les vecteurs qui « rentrent » d'un côté doivent « sortir » de l'autre côté. L'erreur est que le premier vecteur normal est un vecteur qui pointe vers l'intérieur du cylindre, tandis que le second pointe vers l'extérieur. Si nous choisissons, par convention, de prendre uniquement les vecteurs extérieurs, il faut changer le vecteur normal du premier couvercle en  $-r e_x$ . Le premier flux vaudra donc

$$-\pi R^2, \quad (20.310)$$

de telle sorte que le flux total sera nul.  $\triangle$



## 20.13 Formes différentielles exactes et fermées

Nous avons déjà parlé de formes différentielles et de leurs intégrales sur un chemin dans la section 12.24.

DefEFKQmPs

### Définition 20.84.

La forme différentielle  $\omega$  est **exacte** si il existe une fonction  $f$  telle que  $\omega = df$  ; elle est dite **fermée** si  $d\omega = 0$ .

Dire que la forme différentielle  $\omega = fdx + gdy$  est fermée, c'est dire que

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}. \quad (20.311)$$

Il est naturel de se demander si toutes les formes différentielles sont des différentielles de fonctions. Une réponse complète est délicate à établir, mais a d'innombrables conséquences en physique, notamment en ce qui concerne l'existence d'un potentiel vecteur pour le champ magnétique dans les équations de Maxwell.

Le fait qu'une forme exacte soit fermée est relativement facile à établir ; c'est la proposition suivante. La question plus délicate est la réciproque : sous quelles conditions une forme fermée est-elle exacte ?

### Proposition 20.85.

Si  $\omega$  est une 1-forme exacte de classe  $C^1$ , alors  $\omega$  est fermée.

*Démonstration.* Le fait que  $\omega$  soit exacte implique l'existence d'une fonction  $f$  telle que  $\omega = df$ , c'est-à-dire

$$\omega_x = \sum_i a_i(x) dx_i = \sum_i \frac{\partial f}{\partial x_i}(x) dx_i, \quad (20.312)$$

c'est-à-dire que  $a_i(x) = \frac{\partial f}{\partial x_i}(x)$ . L'hypothèse que  $\omega$  est  $C^1$  implique que  $f$  est  $C^2$ , et donc que nous pouvons inverser l'ordre de dérivation pour les dérivées secondes  $\partial_{ij}^2 f = \partial_{ji}^2 f$ . Nous pouvons donc faire le calcul suivant :

$$\frac{\partial a_i}{\partial x_j} = \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} = \frac{\partial a_j}{\partial x_i}, \quad (20.313)$$

ce qu'il fallait démontrer. □

La réciproque est vraie sur un ouvert simplement connexe.

ThoFermeExactFormRappel

### Théorème 20.86.

Supposons que  $D \subset \mathbb{R}^n$  soit un ouvert simplement connexe. Alors toute forme différentielle de degré 1 et de classe  $C^1$  sur  $D$  qui est fermée est exacte.

Nous allons prouver ce théorème dans un cas un peu moins général : celui d'un domaine étoilé de  $\mathbb{R}^2$  plutôt que simplement connexe de  $\mathbb{R}^n$ .

ThoMSofFxL

### Théorème 20.87.

Soit  $D \subset \mathbb{R}^2$ , un ouvert étoilé, et  $\omega$ , une 1-forme fermée de classe  $C^1$ . Alors  $\omega$  est exacte.

*Démonstration.* Soit  $D \subset \mathbb{R}^2$ , un ouvert étoilé par rapport à l'origine. Soient  $f: D \rightarrow \mathbb{R}$ ,  $g: D \rightarrow \mathbb{R}$ , des fonctions de classe  $C^1$  telles que

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x} \quad (20.314)$$

sur  $D$ , et

$$F(x, y) = \int_0^1 [f(tx, ty)x + g(tx, ty)y] dt \quad (20.315)$$

EqIMDefForm133

pour tout  $(x, y) \in D$ .

Étant donné que nous ne définissons  $F(x, y)$  que pour des  $(x, y) \in D$ , la fonction  $t \mapsto f(tx, ty)$  est  $C^1$  sur tout le compact  $[0, 1]$  et aucune divergence de l'intégrale n'est à craindre. Nous sommes donc dans le cadre de la proposition 17.24, et nous pouvons dériver sous le signe intégral.

Nous calculons, en utilisant la règle de dérivation de fonctions composées

$$\begin{aligned} \frac{\partial F}{\partial x}(x, y) &= \int_0^1 \left[ f \frac{\partial f}{\partial x}(tx, ty)x + f(tx, ty) + t \frac{\partial g}{\partial x}(tx, ty)y \right] dt \\ &= \int_0^1 \left[ t \left( x \frac{\partial f}{\partial x}(tx, ty) + y \frac{\partial f}{\partial y}(tx, ty) \right) + f(tx, ty) \right] dt \end{aligned} \quad \text{EqIM133dsdsFlolo} \quad (20.316)$$

où nous avons utilisé l'hypothèse  $\partial_y f = \partial_x g$ . Ce qui se trouve dans la parenthèse n'est autre que  $\partial_t(f(tx, ty))$ , plus précisément, si nous posons  $\mathcal{F}(x, y, t) = f(tx, ty)$ , nous avons

$$\frac{\partial \mathcal{F}}{\partial t}(x, y, t) = x \frac{\partial f}{\partial x}(tx, ty) + y \frac{\partial f}{\partial y}(tx, ty). \quad (20.317)$$

En recopiant le résultat (20.316) en termes de  $\mathcal{F}$ , nous avons

$$\begin{aligned} \frac{\partial F}{\partial x}(x, y) &= \int_0^1 \left( t \frac{\partial \mathcal{F}}{\partial t}(x, y, t) + \mathcal{F}(x, y, t) \right) dt \\ &= \int_0^1 \partial_t(t\mathcal{F}(x, y, t)) dt \\ &= [t\mathcal{F}(x, y, t)]_0^1 \\ &= \mathcal{F}(x, y, 1) \\ &= f(x, y). \end{aligned} \quad (20.318)$$

Le résultat correspondant pour  $\frac{\partial F}{\partial y}(x, y) = g(x, y)$  s'obtient de la même manière. Nous avons donc obtenu que

$$\frac{\partial F}{\partial x} = f, \quad \text{et} \quad \frac{\partial F}{\partial y} = g. \quad \text{EqIMForm133Ffdd} \quad (20.319)$$

En ayant prouvé cela, nous avons prouvé que si  $\omega = f dx + g dy$  avec  $\partial_y f = \partial_x g$ , alors  $\omega = dF$  où  $F$  est définie par (20.315).  $\square$

*Démonstration alternative du théorème 20.87.* Nous posons  $u = tx$  et  $v = ty$ , ainsi que  $\mathcal{F}(x, y, t) = f(u, v)$  et  $\mathcal{G}(x, y, t) = g(u, v)$ . Avec cette notation, nous avons

$$F(x, y) = \int_0^1 (x\mathcal{F}(x, y, t) + y\mathcal{G}(x, y, t)) dt, \quad (20.320)$$

et

$$\begin{aligned} \frac{\partial \mathcal{F}}{\partial x} &= \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} = t \frac{\partial f}{\partial u}, \\ \frac{\partial \mathcal{G}}{\partial x} &= t \frac{\partial g}{\partial u}. \end{aligned} \quad (20.321)$$

Ainsi,

$$\begin{aligned} \frac{\partial F}{\partial x} &= \int_0^1 \left( x \frac{\partial \mathcal{F}}{\partial x} + \mathcal{F} + y \frac{\partial \mathcal{G}}{\partial x} \right) dt \\ &= \int_0^1 \left( xt \frac{\partial f}{\partial u} + \mathcal{F} + yt \frac{\partial g}{\partial u} \right) dt \\ &= \int_0^1 \left[ t \left( x \frac{\partial f}{\partial u} + y \frac{\partial f}{\partial v} \right) + \mathcal{F} \right] dt. \end{aligned} \quad (20.322)$$

où nous avons utilisé le fait que, par hypothèse,  $\frac{\partial g}{\partial u} = \frac{\partial f}{\partial v}$ . Nous calculons par ailleurs que

$$\frac{\partial F}{\partial t} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial t} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial t} = x \frac{\partial f}{\partial u} + y \frac{\partial f}{\partial v}. \quad (20.323)$$

Donc, nous avons

$$\frac{\partial F}{\partial x} = \int_0^1 \left( t \frac{\partial \mathcal{F}}{\partial t} + \mathcal{F} \right) dt = \int_0^1 \frac{\partial}{\partial t} (t\mathcal{F}) dt. \quad (20.324)$$

Par conséquent,

$$\frac{\partial F}{\partial x} = [t\mathcal{F}]_0^1 = \mathcal{F}(x, y, 1) = f(x, y). \quad (20.325)$$

Le même genre de calculs fournit  $\frac{\partial F}{\partial y} = g(x, y)$ .  $\square$

## 20.14 Théorème d'Abel angulaire

ThoTGjmeen

**Théorème 20.88** (Abel angulaire[126]).

Soit  $\sum_n a_n z^n$  une série entière de rayon de convergence plus grand ou égal à 1 et de somme  $f$ . Soit  $\theta_0 \in [0, \frac{\pi}{2}]$ . Nous posons

$$\Delta_{\theta_0} = \{z = 1 - \rho e^{i\varphi} \text{ tel que } \rho > 0, \varphi \in [-\theta_0, \theta_0], |z| < 1\}. \quad (20.326)$$

Nous supposons de plus que  $\sum_n a_n$  converge. Alors

$$\lim_{\substack{z \rightarrow 1 \\ z \in \Delta_0}} f(z) = \sum_{k=0}^{\infty} a_k. \quad (20.327)$$

*Démonstration.* Le résultat de ce théorème est que l'on peut calculer la limite  $z \rightarrow 1$  avec des chemins contenus dans un domaine de la forme de celui dessiné à la figure 20.4.

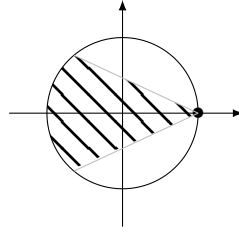


FIGURE 20.4: La zone dans laquelle peut être le chemin qui va vers  $z = 1$ . LabelFigJGuKEjH

De façon très classique nous posons

$$S = \sum_{k=0}^{\infty} a_k \quad S_n = \sum_{k=0}^n a_k, \quad (20.328)$$

et  $R_n = S - S_n$ . En particulier  $a_n = R_{n-1} - R_n$ .

Le but du théorème est de montrer que  $\sum a_n z^n$  converge vers  $S$  lorsque  $z$  converge vers 1 à l'intérieur de  $\Delta_{\theta_0}$ . Pour cela nous calculons pour un  $N$  donné la différence  $\sum_{n=0}^N a_n z^n - S_N$  en triant les termes par ordre de  $R_n$ , en isolant le terme  $R_0$  et le terme  $R_N$  :

$$\sum_{n=0}^N a_n z^n - S_N = \sum_{n=1}^N a_n (z^n - 1) \quad (20.329a)$$

$$= \sum_{n=1}^N (R_{n-1} - R_n)(z^n - 1) \quad (20.329b)$$

$$= R_0(z - 1) + \sum_{n=1}^{N-1} R_n(z^{n+1} - 1 - z^n + 1) + R_N(z^N - 1) \quad (20.329c)$$

$$= R_0(z - 1) + \sum_{n=1}^{N-1} R_n z^n (z - 1) + R_N(z^N - 1) \quad (20.329d)$$

$$= (z - 1) \sum_{n=0}^{N-1} R_n z^n + R_N(z^N - 1). \quad (20.329e)$$

Cela est valable pour tout  $N$  et  $|z| < 1$ . Nous avons donc

$$\sum_{n=0}^N a_n z^n - S_N = (z-1) \sum_{n=0}^{N-1} R_n z^n + R_N (z^N - 1). \quad (20.330)$$

Par hypothèse nous avons  $\lim_{N \rightarrow \infty} R_N = 0$ . Et de plus le membre de gauche converge parce que chacun des deux termes converge séparément. En passant à la limite nous avons pour tout  $|z| < 1$  :

$$f(z) - S = (z-1) \sum_{n=0}^{\infty} R_n z^n. \quad (20.331)$$

Nous voudrions étudier le comportement de la différence  $f(z) - S$  lorsque  $z$  tend vers 1. Pour cela nous nous fixons  $\epsilon > 0$  et  $N \geq 1$  tel que  $|R_n| < \epsilon$  dès que  $n \geq N$ . Alors pour tout  $|z| < 1$  nous avons

$$|f(z) - S| \leq |z-1| \left( \sum_{n=0}^N |R_n| \underbrace{|z^n|}_{\leq 1} + \sum_{n=N+1}^{\infty} \underbrace{|R_n|}_{\leq \epsilon} |z^n| \right) \quad (20.332a)$$

$$\leq |z-1| \sum_{n=0}^N |R_n| + \epsilon \frac{|z-1|}{1-|z|} \quad (20.332b)$$

où nous avons utilisé la somme de la série géométrique (1.840) et l'égalité  $|z^n| = |z|^n$ . Avant de nous particulariser à  $z \in \Delta_{\theta_0}$  nous devons anticiper un problème au dénominateur en multipliant par le binôme conjugué :

$$\frac{|z-1|}{1-|z|} = \frac{|z-1|(1+|z|)}{1-|z|^2}. \quad (20.333)$$

C'est maintenant que nous nous particularisons à  $z \in \Delta_{\theta_0}$  en posant  $z = \rho e^{i\varphi}$  et en remarquant que  $|z|^2 = 1 - 2\rho \cos(\varphi) + \rho^2$ . Nous avons le calcul suivant :

$$\frac{|z-1|}{1-|z|} = \frac{\rho(1+|z|)}{2\rho \cos(\varphi) - \rho^2} \quad (20.334a)$$

$$= \frac{1+|z|}{2 \cos(\varphi) - \rho} \quad (20.334b)$$

$$\leq \frac{2}{2 \cos(\varphi) - \rho} \quad (20.334c)$$

$$\leq \frac{2}{2 \cos(\varphi) - \cos(\theta_0)} \quad (20.334d)$$

$$\leq \frac{2}{2 \cos(\theta_0) - \cos(\theta_0)} \quad (20.334e)$$

$$= \frac{2}{\cos(\theta_0)}. \quad (20.334f)$$

Quelques justifications.

- Vu que nous avons dans l'idée de faire  $\rho \rightarrow 0$  nous supposons que  $\rho < \cos(\theta_0)$ .
- Nous avons  $\cos(\varphi) > \cos(\theta_0)$  parce que  $z$  est dans  $\Delta_{\theta_0}$ .

Nous avons donc, pour tout  $z \in \Delta_{\theta_0}$  que

$$|f(z) - S| \leq |z-1| \sum_{n=0}^N |R_n| + \epsilon \frac{2}{\cos(\theta_0)}. \quad (20.335)$$

Il suffit de prendre  $\rho$  assez petit pour que

$$|z-1| \sum_{n=0}^N |R_n| < \epsilon \quad (20.336)$$

et nous avons

$$|f(z) - S| \leq \epsilon \left( 1 + \frac{2}{\cos(\theta_0)} \right). \quad (20.337)$$

Nous avons donc bien  $\lim_{\substack{z \rightarrow 1 \\ z \in \Delta_0}} f(z) = S$ , comme nous le voulions.  $\square$

La réciproque du théorème d'Abel angulaire est que si  $f(z) = \sum_n a_n z^n$  sur  $B(0, 1)$  se prolonge par continuité en  $z = 1$  alors cette prolongation se fait par  $f(1) = \sum_n a_n$ . Cela est faux comme le montre l'exemple suivant.

**Exemple 20.89.**

Nous considérons la série entière  $\sum_{n=0}^{\infty} (-1)^n z^n$  qui converge<sup>20</sup> vers

$$f(z) = \frac{1}{1+z} \quad (20.338)$$

sur  $B(0, 1)$ . De plus nous avons

$$\lim_{\substack{z \rightarrow 1 \\ |z| < 1}} \frac{1}{1+z} = \frac{1}{2}. \quad (20.339)$$

Donc la fonction converge bien vers quelque chose lorsque  $z$  tend vers 1. La fonction  $f$  se prolonge par continuité en 1. Pourtant la série ses coefficients  $\sum_n (-1)^n$  ne converge pas.  $\triangle$

Le théorème suivant donne une espèce d'inverse au théorème d'Abel angulaire. En effet il dit que si la série converge en allant vers 1 le long de l'axe réel, alors ça converge vers la somme des coefficients. Il faut cependant une hypothèse en plus sur les  $a_n$ .

**Théorème 20.90** (Théorème taubérien faible[126]).

Soit  $\sum_n a_n z^n$  une série entière de rayon de convergence 1 et de somme  $f$ . Nous supposons

(1) Il existe  $S \in \mathbb{C}$  tel que  $\lim_{\substack{x \rightarrow 1 \\ x \in ]-1, 1[}} f(x) = S$ .

(2)  $\lim_{n \rightarrow \infty} n a_n = 0$ .

Alors la série  $\sum_{n=0}^{\infty} a_n$  converge et vaut  $S$ .

*Démonstration.* Nous notons  $S_n = \sum_{k=0}^n a_k$  et  $M = \sup_{k \geq 1} k|a_k|$ , qui est fini par hypothèse. Pour  $x \in ]0, 1[$  et  $n \geq 0$  nous avons

$$S_n - f(x) = \sum_{k=1}^n a_k - \sum_{k=1}^n a_k x^k - \sum_{k=n+1}^{\infty} a_k x^k = \sum_{k=1}^n a_k (1 - x^k) - \sum_{k=n+1}^{\infty} a_k x^k. \quad (20.340)$$

Nous utilisons la série géométrique sous la forme  $1 - x^k = (1 - x) \sum_{i=0}^{k-1} x^i$  pour écrire

$$S_n - f(x) = \sum_{k=1}^n a_k (1 - x) \underbrace{\sum_{i=0}^{k-1} x^i}_{\leq k} - \sum_{k=n+1}^{\infty} a_k x^k \quad (20.341a)$$

$$\leq \sum_{k=1}^n k a_k (1 - x) - \sum_{k=n+1}^{\infty} a_k x^k, \quad (20.341b)$$

---

20. C'est la série géométrique de raison  $-z$ .

donc en passant à la norme

$$|S_n - f(x)| \leq (1-x)Mn + \sum_{k=n+1}^{\infty} |a_k| x^k \quad (20.342a)$$

$$\leq (1-x)Mn + \sum_{k=n+1}^{\infty} \underbrace{\frac{k}{n} |a_k|}_{\leq M/n} x^k \quad (20.342b)$$

$$\leq (1-x)Mn + \frac{M}{n} \sum_{k=n+1}^{\infty} x^k \quad (20.342c)$$

$$\leq (1-x)Mn + \frac{M}{n} \frac{1}{1-x}. \quad (20.342d)$$

Ce que nous cherchons à étudier est le comportement  $x \rightarrow 1$  et montrer que  $S_n \rightarrow S$ , ce qui nous incite à calculer  $|S_n - f(1 - \frac{\epsilon}{n})|$  avec  $0 < \epsilon < 1$  :

$$|S_n - f(1 - \frac{\epsilon}{n})| \leq \epsilon M + \epsilon. \quad (20.343)$$

Nous choisissons  $N_1$  tel que  $\frac{M}{n} \leq \epsilon^2$  dès que  $n \geq N_1$ . En sus nous savons que

$$\lim_{\epsilon \rightarrow 0} f(1 - \epsilon) = S. \quad (20.344)$$

Nous choisissons  $N_2$  de telle sorte à avoir

$$\left| f\left(1 - \frac{\epsilon}{n}\right) - S \right| < \epsilon, \quad (20.345)$$

et en prenant  $n \geq \max(N_1, N_2)$  nous avons

$$|S_n - S| \leq \left| S_n - f\left(1 - \frac{\epsilon}{n}\right) \right| + \left| f\left(1 - \frac{\epsilon}{n}\right) - S \right| \leq \epsilon M + 2\epsilon. \quad (20.346)$$

Il suffit de choisir  $\epsilon$  suffisamment petit (en particulier pour que  $\epsilon M$  soit petit) pour montrer que  $|S_n - S|$  est borné par un nombre arbitrairement petit.  $\square$

## 20.15 Passage à la limite sous le signe intégral

Un autre résultat très important pour l'étude de l'intégrabilité est le théorème de la **convergence dominée de Lebesgue** :

### Théorème 20.91.

Soit  $E \subset \mathbb{R}^n$  un ensemble mesurable et  $\{f_k\}$ , une suite de fonctions intégrables sur  $E$  qui converge simplement vers une fonction  $f: E \rightarrow \mathbb{R}$ . Supposons qu'il existe une fonction  $g$  intégrable sur  $E$  telle que pour tout  $k$ ,

$$|f(x)| \leq g(x) \quad (20.347)$$

pour tout  $x \in E$ . Alors  $f$  est intégrable sur  $E$  et

$$\int_E f = \lim_{k \rightarrow \infty} \int_E f_k. \quad (20.348)$$

### 20.15.1 Intégrale en dimension un

#### Proposition 20.92 (Critère de comparaison).

Soit  $f$  mesurable sur  $]a, \infty[$  et bornée sur tout  $]a, b]$ , et supposons qu'il existe un  $X_0 \geq a$ , tel que sur  $]X_0, \infty[$ ,

$$|f(x)| \leq g(x) \quad (20.349)$$

où  $g(x)$  est intégrable. Alors  $f(x)$  est intégrable sur  $]a, \infty[$ .

**Corolaire 20.93** (Critère d'équivalence).

Soient  $f$  et  $g$  des fonctions mesurables et positives ou nulles sur  $]a, \infty[$ , bornées sur tout  $]a, b]$ , telles que

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = L \quad (20.350)$$

existe dans  $\bar{\mathbb{R}}$ .

- (1) Si  $L \neq \infty$  et  $\int_a^\infty g(x)dx$  existe, alors  $\int_a^\infty f(x)dx$  existe,
- (2) Si  $L \neq 0$  et si  $\int_a^\infty f(x)dx$  existe, alors  $\int_a^\infty g(x)dx$  existe,

CorCritFonsTest

**Corolaire 20.94** (Critère des fonctions test).

Soit  $f(x)$  une fonction mesurable et positive ou nulle sur  $]a, \infty[$  et bornée pour tout  $]a, b]$ . Nous posons

$$L(\alpha) = \lim_{x \rightarrow \infty} x^\alpha f(x), \quad (20.351)$$

et nous supposons qu'elle existe.

- (1) Si il existe  $\alpha > 1$  tel que  $L(\alpha) \neq \infty$ , alors  $\int_a^\infty f(x)dx$  existe,
- (2) Si il existe  $\alpha \leq 1$  et  $L(\alpha) \neq 0$ , alors  $\int_a^\infty f(x)dx$  n'existe pas.

CorAlphaLCasInteabf

**Corolaire 20.95.**

Soit  $f: ]a, b] \rightarrow \mathbb{R}$  une fonction mesurable, positive ou nulle, et bornée sur  $[a + \epsilon, b]$   $\forall \epsilon > 0$ . Si  $\lim_{x \rightarrow a} (x - a)^\alpha f(x) = L$  existe, alors

- (1) Si  $\alpha < 1$  et  $L \neq \infty$ , alors  $\int_a^b f(x)dx$  existe,
- (2) Si  $\alpha \geq 1$  et  $L \neq 0$ , alors  $\int_a^b f(x)dx$  n'existe pas.

## 20.15.2 Intégrales convergentes

**Définition 20.96.**

Soit  $f$ , une fonction mesurable sur  $[a, \infty[$ , bornée sur tout intervalle  $[a, b]$ . On dit que l'intégrale

$$\int_a^\infty f(x)dx \quad (20.352)$$

converge si la limite

$$\lim_{X \rightarrow \infty} \int_a^X f \quad (20.353)$$

EqDefConvergeZeroInfX

existe et est finie.

## 20.15.3 La méthode de Rothstein-Trager

subSecBCRYooRVjFpS

Mes sources pour parler d'intégration de fractions rationnelles : [655].

ThoXJFatfu

**Théorème 20.97** (Rothstein-Trager[656]).

Soient  $P, Q \in \mathbb{Q}[X]$  premiers entre eux avec  $\text{pgcd}(P, Q) = 1$  et  $\deg(P) < \deg(Q)$ . Nous supposons que  $Q$  est unitaire et sans facteurs carrés. Supposons que nous puissions écrire, dans une extension  $\mathbb{K}$  de  $\mathbb{Q}$  la primitive de  $P/Q$  de la façon suivante :

$$\int \frac{P}{Q} = \sum_{i=1}^n c_i \ln(P_i) \quad (20.354)$$

EqCHVaDay

où les  $c_i$  sont des constantes non nulles et deux à deux distinctes et où les  $P_i$  sont des polynômes unitaires non constants sans facteurs carrés et premiers deux à deux entre eux dans  $\mathbb{K}[X]$ .

Alors les  $c_i$  sont les racines distinctes du polynôme

$$R(Y) = \text{res}_X(P - YQ', Q) \in \mathbb{K}[Y] \quad (20.355)$$

et

$$P_i = \text{pgcd}(P - c_i Q', Q). \quad (20.356)$$

*Démonstration.* Nous posons

$$U_i = \prod_{j \neq i} P_j. \quad (20.357)$$

- (1) **Question de division** Ensuite nous dérivons formellement l'équation (20.354) et nous multiplions les deux côtés du résultat par  $\prod_{j=1}^n P_j$  :

$$P \prod_{j=1}^n P_j = Q \sum_{i=1}^n c_i \frac{P'_i}{P_i} \prod_{j=1}^n P_j = Q \sum_{i=1}^n c_i P'_i U_i. \quad (20.358)$$

Une première chose que nous en tirons est que  $Q$  divise le produit  $P \prod_{j=1}^n P_j$  ; mais  $P$  et  $Q$  étant premiers entre eux,

$$Q \mid \prod_{j=1}^n P_j \quad (20.359)$$

par le théorème de Gauss 6.50.

Une seconde chose que nous tirons de (20.358) est que  $P_j$  divise  $Q \sum_{i=1}^n c_i P'_i U_i$ . De cette somme, à cause du  $U_i$  qui est divisé par  $P_j$  pour tout  $i$  sauf  $i = j$ , le polynôme  $P_j$  divise tous les termes sauf peut-être un. Donc il les divise tous et en particulier

$$P_j \mid Q c_j P'_j U_j \quad (20.360)$$

En nous souvenant que les  $P_k$  sont premiers entre eux,  $P_j$  ne divise pas  $U_j$ . De plus  $P_j$  étant sans facteurs carrés, les polynômes  $P_j$  et  $P'_j$  sont premiers entre eux. Il ne reste que  $Q$ . Nous en déduisons que

$$P_j \mid Q \quad (20.361)$$

pour tout  $1 \leq j \leq n$ . Et vu que les  $P_i$  sont premiers entre eux, le fait que chacun divise  $Q$  implique que leur produit divise  $Q$ , c'est-à-dire

$$\prod_{j=1}^n P_j \mid Q. \quad (20.362)$$

Or nous avons déjà prouvé la division contraire. Du fait que les deux polynômes sont unitaires nous en déduisons qu'ils sont en réalité égaux :

$$Q = \prod_{j=1}^n P_j. \quad (20.363)$$

Nous pouvons simplifier les deux membres de (20.358) par cela :

$$P = \sum_{i=1}^n c_i P'_i U_i. \quad (20.364)$$

- (2) **Encore un peu de division**

En dérivant (20.363) nous trouvons

$$Q' = \sum_{j=1}^n P'_j U_j, \quad (20.365)$$

et en écrivant  $P$  sous sa forme (20.364),

$$P - c_i Q' = \sum_{j=1}^n c_j P'_j U_j - \sum_{j=1}^n c_i P'_j U_j = \sum_{j=1}^n (c_j - c_i) P'_j U_j. \quad (20.366)$$

Le terme  $i = j$  de la somme est nul ; en ce qui concerne les autres termes, ils sont divisés par  $P_i$  parce que  $P_i \mid U_j$ . Donc  $P_i$  divise tous les termes de la somme et nous avons

$$P_i \mid P - c_i Q'. \quad (20.367)$$



(3) **Un pgcd pour continuer**

Nous montrons à présent que  $P_i = \text{pgcd}(P - c_i Q', Q)$ . Pour cela nous utilisons la multiplicité du PGCD lorsque les facteurs sont premiers entre eux :

$$\text{pgcd}(P - c_i Q', Q) = \text{pgcd}(P - c_i Q', \prod_{j=1}^n P_j) = \prod_{j=1}^n \text{pgcd}(P - c_i Q', P_j). \quad (20.368)$$

Nous remplaçons  $P - c_i Q'$  par son expression (20.366) et nous écrivons un des facteurs du produit :

$$\text{pgcd}(P - c_i Q', P_j) = \text{pgcd}\left(\sum_{k=1}^n (c_k - c_i) P'_k U_k, P_j\right) \quad (20.369)$$

Le polynôme  $P_j$  divise tous les  $U_k$  sauf celui avec  $k = j$ . Donc le lemme 6.57(1) nous permet de dire

$$\text{pgcd}(P - c_i Q', P_j) = \text{pgcd}((c_j - c_i) P'_j U_j, P_j) = \begin{cases} 1 & \text{si } i \neq j \\ P_j & \text{si } i = j. \end{cases} \quad (20.370)$$

La seconde ligne provient du fait que nous avons déjà montré que  $P_j \mid P - c_j Q'$ . En fin de compte,

$$\text{pgcd}(P - c_i Q', Q) = P_i. \quad (20.371)$$

(4) **Une histoire de résultant**

Les nombres  $c_i$  sont tels que les polynômes  $P - c_i Q'$  et  $Q$  ne sont pas premiers entre eux. Vu que les  $P_i$  sont non nuls, la proposition 8.24 nous dit que le résultant

$$\text{res}_X(P - c_i Q', Q) = 0. \quad (20.372)$$

Donc les  $c_i$  sont des racines du polynôme (en  $Y$ )

$$R(Y) = \text{res}_X(P - Y Q', Q). \quad (20.373)$$

Nous n'avons pas prouvé qu'ils étaient *toutes* les racines<sup>21</sup>.

(5) **Toutes les racines**

Nous allons maintenant montrer que les  $c_i$  étaient toutes les racines imaginables du polynôme (20.373) dans toutes les extensions de  $\mathbb{Q}$ . Soit donc  $c$  une racine de  $R$  dans une extension  $\hat{\mathbb{K}}$  de  $\mathbb{K}$  qui ne soit pas parmi les  $c_i$  de la formule (20.354). Étant donné que  $c$  est racine du résultant, les polynômes  $P - c Q'$  et  $Q$  ont un PGCD non trivial, c'est-à-dire non constant. Donc

$$\text{pgcd}(P - c Q', Q) = S \in \hat{\mathbb{K}}[X] \quad (20.374)$$

est un polynôme non constant. Si  $T$  un facteur irréductible de  $S$ , alors  $T$  divise  $P - c Q'$  et  $Q$ , mais  $Q = \prod_{i=1}^n P_i$  avec les  $P_i$  premiers entre eux. Donc  $T$  ne peut diviser que l'un (et exactement un) d'entre eux<sup>22</sup>. Soit  $P_{i_0}$  celui qui est divisé par  $T$ . La relation (20.366) dans ce contexte donne :

$$P - c Q' = \sum_{j=1}^n (c_j - c) P'_j U_j \quad (20.375)$$

Le polynôme  $T$  divise tous les  $U_j$  avec  $j \neq i_0$ , mais comme en plus il divise  $P - c Q'$ , il divise aussi le dernier terme de la somme :

$$T \mid (c_{i_0} - c) P'_{i_0} U_{i_0}. \quad (20.376)$$

Le polynôme  $T$  ne divisant pas  $U_{i_0}$  et  $(c_{i_0} - c)$  étant non nul, nous concluons que  $T$  divise  $P'_{i_0}$ . Mais cela n'est pas possible parce que nous avons supposé que  $P_{i_0}$  était sans facteur carré, ce qui voulait entre autres dire que  $P_{i_0}$  et  $P'_{i_0}$  n'ont pas de facteurs communs.

21. De plus, nous n'avons pas de garanties que ces racines soient dans  $\mathbb{Q}$ , et en fait il y a des cas dans lesquels les  $c_i$  n'y sont pas.

22. On ne peut pas diviser deux trucs qui sont premiers entre eux ; c'est une question de cohérence, madame !

□

Ce théorème suggère la méthode suivante pour trouver la primitive de la fraction rationnelle  $P/Q$  (si elle vérifie les hypothèses)

- (1) Écrire le résultant  $R(y) = \text{res}_X(P - yQ', Q)$  et en trouver les racines  $\{c_i\}_{i=1,\dots,n}$ .
- (2) Calculer les polynômes  $P_i = \text{pgcd}(P - c_iQ', Q)$ .
- (3) Écrire la réponse :

$$\int \frac{P}{Q} = \sum_{i=1}^n c_i \ln(P_i). \quad (20.377)$$

Notons que le polynôme  $R(Y)$  est de degré  $\deg(Q)$  (pour le voir, faire un peu de comptage de lignes et colonnes dans la matrice de Sylvester), donc il n'est à priori pas pire à factoriser que  $Q$  lui-même<sup>23</sup>. Mais il se peut que nous ayons de la chance et que  $R$  soit plus facile que  $Q$ .

À part qu'on a peut-être plus de chance avec  $R$  qu'avec  $Q$ , l'avantage de la méthode est qu'elle permet d'éviter de passer par des extensions de  $\mathbb{Q}$  non nécessaires<sup>24</sup>.

**Exemple 20.98** ([656]).

Prenons la fraction rationnelle  $\frac{x}{x^2-3}$ . L'intégration via les fractions simples est :

$$\int \frac{x}{x^2-3} dx = \frac{1}{2} \int \frac{1}{x-\sqrt{3}} + \frac{1}{2} \int \frac{1}{x+\sqrt{3}} = \frac{1}{2} \ln(x-\sqrt{3}) + \frac{1}{2} \ln(x+\sqrt{3}) = \frac{1}{2} \ln(x^2-3). \quad (20.378)$$

Nous voyons que dans la réponse, il n'y a pas de racines. Passer par l'extension  $\mathbb{Q}[\sqrt{3}]$  est par conséquent peut-être un effort inutile. Voyons comment les choses se mettent avec la méthode Rothstein-Trager.

D'abord

$$R(Y) = \text{res}_X(X - 2YX, X^2 - 3) \quad (20.379a)$$

$$= \text{res}_X((1 - 2Y)X, X^2 - 3) \quad (20.379b)$$

$$= \det \begin{pmatrix} 1 - 2Y & 0 & 0 \\ 0 & 1 - 2Y & 0 \\ 1 & 0 & -3 \end{pmatrix} \quad (20.379c)$$

$$= (1 - 2Y)(-3(1 - 2Y)) \quad (20.379d)$$

$$= -3(1 - 2Y)^2, \quad (20.379e)$$

dont les solutions sont faciles : il n'y a que la racine double  $y = \frac{1}{2}$ . La somme (20.354) sera donc réduite à un seul terme avec  $c_1 = \frac{1}{2}$ . Nous calculons  $P_1$  :

$$P_1 = \text{pgcd}(X - \frac{1}{2}2X, X^2 - 3) = \text{pgcd}(0, X^2 - 3) = X^2 - 3, \quad (20.380)$$

et par conséquent

$$\int \frac{X}{X^2-3} = \frac{1}{2} \ln(X^2-3). \quad (20.381)$$

À aucun moment nous ne sommes sortis de  $\mathbb{Q}$ . △

Comme vu sur cet exemple, l'intérêt du théorème de Rothstein-Trager est de permettre, lorsqu'on a de la chance, d'en profiter, et non de nous en rendre compte à la fin en remarquant bêtement que la réponse pouvait s'écrire dans  $\mathbb{Q}[X]$ .

23. C'est de la factorisation de  $Q$  qu'on a besoin pour utiliser la méthode de décomposition en fractions simples.

24. J'imagine que pour un ordinateur, c'est plus facile d'éviter les extensions.

**Remarque 20.99.**

Afin d'utiliser cette méthode, il faut s'assurer que  $Q$  soit sans facteurs carrés. Si nous devons intégrer un  $\frac{P}{Q}$  quelconque, nous devons commencer par écrire

$$Q = Q_1 Q_2^2 Q_3^3 \dots Q_r^r, \quad (20.382)$$

et ensuite il y a moyen de ramener l'intégrale de  $P/Q$  à des intégrales de  $\frac{P}{Q_1 \dots Q_r}$ . Cela ne demande pas de factoriser complètement  $Q$ , mais seulement de trouver ses facteurs irréductibles  $Q_i$  dans  $\mathbb{Q}[X]$ .

Dans l'exemple donné plus haut,  $Q = X^2 - 3$  a des facteurs irréductibles autres que  $Q$  lui-même dans  $\mathbb{R}[X]$ , mais nous n'en avons pas besoin.

Voici une exemple où nous évitons de passer par les complexes.

EXooIPEQooGKDjea

**Exemple 20.100** ([657]).

À calculer :  $\int \frac{1}{x^3+x}$ . La décomposition en fractions simples donne :

$$\frac{1}{x^3+x} = \frac{1}{x} - \frac{1/2}{x-i} - \frac{1/2}{x+i}. \quad (20.383)$$

Déjà cette décomposition passe par l'extension  $\mathbb{Q}[i]$ , et le calcul de la primitive de  $\frac{1}{x+i}$  demande le logarithme complexe qui ne sera vu que dans la proposition 26.43 au prix d'un peu de sang. Nous verrons dans l'exemple 26.44 comment ça se passe en passant par les complexes.

En ce qui concerne la méthode de Rothstein-Trager, nous commençons par calculer le résultant (qui est tout de même un peu de calcul) :

$$P(y) = \text{res}_X (-3yX^2 - y + 1, X^3 - X) \quad (20.384a)$$

$$= \det \begin{pmatrix} -3y & 0 & 1-y & 0 & 0 \\ 0 & -3y & 0 & 1-y & 0 \\ 0 & 0 & -3y & 0 & 1-y \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} \quad (20.384b)$$

$$= -(y-1)^2(2y+1)^2 \quad (20.384c)$$

```
-----
| Sage Version 5.7, Release Date: 2013-02-19                |
| Type "notebook()" for the browser-based notebook interface. |
| Type "help()" for help.                                   |
|-----|
```

```
sage: y=var('y')
sage: R=matrix(5,5,[-3*y,0,1-y,0,0,0,-3*y,0,1-y,0,0,0,
                    -3*y,0,1-y,1,0,1,0,0,0,1,0,1,0])
sage: R.determinant().factor()
-(y - 1)*(2*y + 1)^2
```

Les solutions sont  $c_1 = 1$  et  $c_2 = -\frac{1}{2}$ . Nous pouvons alors calculer les  $P_i$  :

$$P_1 = \text{pgcd}(-3X^2, X^3 + X) = X \quad (20.385)$$

et

$$P_2 = \text{pgcd}\left(\frac{3}{2}X^2 + \frac{3}{2}, X^3 + X\right) = X^2 + 1, \quad (20.386)$$

et finalement

$$\int \frac{1}{X^3 + X} = \ln(X) - \frac{1}{2} \ln(X^2 + 1). \quad (20.387)$$

△

Notons qu'il n'y a pas de miracles : lorsque la réponse contient des racines, nous ne pouvons pas couper à passer par des extensions et factoriser un peu à la dure.

ExYQODuyU

**Exemple 20.101** ([657]).

Nous voulons calculer

$$\int \frac{1}{X^2 + 1}. \quad (20.388)$$

Nous posons donc  $P = 1$  et  $Q = X^2 + 1$ . Le résultant à calculer est

$$P(y) = \text{res}_X(-2yX + 1, X^2 + 1) = \det \begin{pmatrix} -2y & 1 & 0 \\ 0 & -2y & 1 \\ 1 & 0 & 1 \end{pmatrix} = 4y^2 + 1. \quad (20.389a)$$

Les racines de cela sont complexes et il n'y a donc pas d'échappatoires :  $c_1 = \frac{i}{2}$ ,  $c_2 = -\frac{i}{2}$ . Ensuite, étant donné que  $X^2 + 1 = (X + i)(X - i) = i(-iX + 1)(X - i)$  nous avons

$$P_1 = \text{pgcd}(-iX + 1, X^2 + 1) = X + i. \quad (20.390)$$

Notons que de façon naturelle, nous aurions écrit  $P_1 = 1 - iX$ , mais par convention nous considérons le PGCD unitaire. Cela ne change rien à la réponse parce que changer  $P_i$  en  $kP_i$  ne fait que rajouter une constante  $\ln(k)$  à la primitive trouvée.

De la même façon,

$$P_2 = \text{pgcd}(1 + iX, X^2 + 1) = X - i. \quad (20.391)$$

Au final nous écrivons

$$\int \frac{1}{X^2 + 1} = \frac{i}{2} \ln(X + i) - \frac{i}{2} \ln(X - i). \quad (20.392)$$

△

**Remarque 20.102.**

Tout cela est si nous voulons absolument écrire la primitive avec des logarithmes de polynômes. Pour celui de l'exemple 20.101, nous avons trouvé

$$\int \frac{1}{X^2 + 1} = \frac{i}{2} \ln(X + i) - \frac{i}{2} \ln(X - i). \quad (20.393)$$

EaBCiCCbs

Mais

```
sage: f(x)=1/(x**2+1)
sage: f.integrate(x)
x |--> arctan(x)
```

Si nous acceptons de passer aux fonctions trigonométriques (inverses), la primitive prend un tour très différent et bien réel. Ces deux visions de l'univers sont bien entendu <sup>25</sup> compatibles. En effet, afin de tomber juste, nous allons prendre la primitive

$$f(x) = \frac{i}{2} \ln(ix - 1) - \frac{i}{2} \ln(ix + 1) \quad (20.394)$$

au lieu de (20.393). Il s'agit seulement de multiplier l'intérieur des logarithmes, ce qui ne donne qu'une constante de différence. Ensuite nous passons à la forme trigonométrique des nombres complexes :  $ix - 1 = \sqrt{x^2 + 1}e^{i \arctan(-x)}$  et  $ix + 1 = \sqrt{x^2 + 1}e^{i \arctan(x)}$ . Avec un peu de calcul,

$$f(x) = -\frac{1}{2} \left( \arctan(-x) - \arctan(x) \right) = \arctan(x). \quad (20.395)$$

---

25. Si on croit que la mathématique est cohérente.

## 20.16 Rappel sur les intégrales usuelles

Soit une fonction

$$\begin{aligned} f: [a, b] \subset \mathbb{R} &\rightarrow \mathbb{R}^+ \\ x &\mapsto f(x). \end{aligned} \quad (20.396)$$

L'intégrale de  $f$  sur le segment  $[a, b]$ , notée  $\int_a^b f(x)dx$  est le nombre égal à l'aire de la surface située entre le graphe de  $f$  et l'axe des  $x$ , comme indiqué à la figure 20.5.

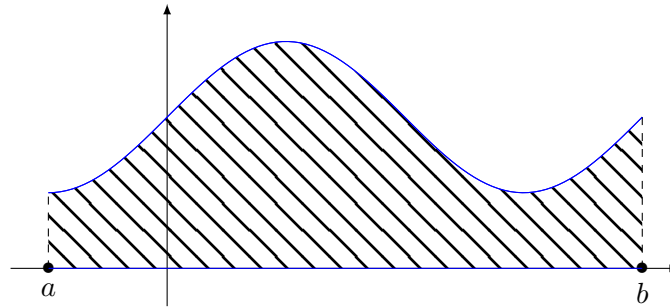


FIGURE 20.5: L'intégrale de  $f$  entre  $a$  et  $b$  représente la surface sous la fonction. Label: gKfKLooMbJxdI

### Définition 20.103.

Si  $f$  est une fonction de une variable à valeurs réelles, une **primitive** de  $f$  est une fonction  $F$  telle que  $F' = f$ .

Toute fonction continue admet une primitive.

### Théorème 20.104 (Théorème fondamental du calcul intégral).

Si  $f$  est une fonction positive et continue, et si  $F$  est une primitive de  $f$ , alors

$$\int_a^b f(x)dx = F(b) - F(a). \quad (20.397)$$

### Remarque 20.105.

Si  $f$  est une fonction continue par morceaux, l'intégrale de  $f$  se calcule comme la somme des intégrales de ses morceaux. Plus précisément si nous avons  $a = x_0 < x_1 < \dots < x_n = b$  et si  $f$  est continue sur  $]x_i, x_{i+1}[$  pour tout  $i$ , alors nous posons

$$\int_a^b f(x)dx = \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \dots + \int_{x_{n-1}}^{x_n} f(x)dx. \quad (20.398)$$

Sur chacun des morceaux, l'intégrale se calcule normalement en passant par une primitive.

## 20.17 Intégrales le long de chemins

### 20.17.1 Circulation d'un champ de vecteur

#### Définition 20.106.

Soit  $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  un champ de vecteurs et un chemin  $\sigma: [a, b] \rightarrow \mathbb{R}^3$ . On appelle **circulation** de  $F$  le long du chemin  $\sigma$  le scalaire

$$\int_a^b F(\sigma(t)) \cdot \sigma'(t)dt. \quad (20.399) \quad \text{EqDefvkZwh}$$

Il existe de nombreuses notations pour cela ; entre autres :

$$\int_{\sigma} F = \int_{\sigma} F \cdot ds. \quad (20.400)$$

En physique, la circulation de la force le long d'un chemin est le travail de la force.

**Exemple 20.107.**

À la surface de la Terre, le champ de gravitation est donné par

$$G(x, y, z) = -mg \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \quad (20.401)$$

Si nous considérons un mobile qui monte à vitesse constante jusqu'à la hauteur  $h$ , c'est-à-dire le chemin

$$\sigma(t) = \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix} \quad (20.402)$$

avec  $t \in [0, h]$ . Le travail de la gravitation est alors donné par

$$W = \int_0^h G(\sigma(t)) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} dt = -mg \int_0^h \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} dt = -mgh. \quad (20.403)$$

Cela est bien le résultat usuel de l'énergie potentielle. Nous allons voir bientôt que nous nommons la fonction  $mgh$  énergie *potentielle* précisément parce que la force dérive de ce potentiel.  $\triangle$

**Exemple 20.108.**

Soit le chemin

$$\begin{aligned} \sigma: [0, 2\pi] &\rightarrow \mathbb{R}^3 \\ t &\mapsto \begin{pmatrix} \sin(t) \\ \cos(t) \\ t \end{pmatrix}. \end{aligned} \quad (20.404)$$

et le champ de vecteurs

$$F \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (20.405)$$

La circulation de ce champ de vecteur le long de l'hélice  $\sigma$  est

$$\begin{aligned} \int_{\sigma} F \cdot ds &= \int_0^{2\pi} (F \circ \sigma)(t) \cdot \sigma'(t) dt \\ &= \int_0^{2\pi} \begin{pmatrix} \sin(t) \\ \cos(t) \\ t \end{pmatrix} \cdot \begin{pmatrix} \cos(t) \\ -\sin(t) \\ 1 \end{pmatrix} dt \\ &= \int_0^{2\pi} t dt \\ &= \left[ \frac{t^2}{2} \right]_0^{2\pi} \\ &= 2\pi^2. \end{aligned} \quad (20.406)$$

$\triangle$

**Proposition 20.109.**

La circulation d'un champ de vecteurs le long d'un chemin ne dépend pas du paramétrage. En d'autres termes, si  $\sigma_1$  et  $\sigma_2$  sont deux chemins équivalents, alors

$$\int_{\sigma_1} F = \int_{\sigma_2} F. \quad (20.407)$$

*Démonstration.* Soient deux chemins  $\sigma_1: [a, b] \rightarrow \mathbb{R}^3$  et  $\sigma_2: [c, d] \rightarrow \mathbb{R}^3$  équivalents, c'est-à-dire tels que

$$\sigma_1(t) = \sigma_2(\varphi(t)) \quad (20.408)$$

où  $\varphi: [a, b] \rightarrow [c, d]$  strictement croissante. En utilisant le fait que  $\sigma_1'(t) = \varphi'(t)\sigma_2'(\varphi(t))$ , nous avons

$$\begin{aligned} \int_{\sigma_1} F \cdot ds &= \int_a^b F(\sigma_1(t)) \cdot \sigma_1'(t) dt \\ &= \int_a^b F(\sigma_2(\varphi(t))) \cdot \sigma_2'(\varphi(t)) \varphi'(t) dt \\ &= \int_{\varphi(a)}^{\varphi(b)} F(\sigma_2(s)) \cdot \sigma_2'(s) ds \\ &= \int_c^d F(\sigma_2(s)) \cdot \sigma_2'(s) ds \\ &= \int_{\sigma_2} F \cdot ds. \end{aligned} \quad (20.409)$$

où nous avons effectué le changement de variables  $s = \varphi(t)$ ,  $ds = \varphi'(t)dt$ .  $\square$

### Remarque 20.110.

Si  $\sigma_2$  est le chemin opposé de  $\sigma$ , alors

$$\int_{\sigma_2} F = - \int_{\sigma_1} F. \quad (20.410)$$

## 20.18 Circulation d'un champ conservatif

Si nous avons une fonction scalaire  $V: \mathbb{R}^3 \rightarrow \mathbb{R}$ , nous pouvons construire un champ de vecteur en prenant le gradient :

$$F(x) = \nabla V(x). \quad (20.411)$$

On dit que le champ de vecteur  $F$  **dérive** de  $V$ , et on dit que  $V$  est le **potentiel** de  $F$ . Nous posons la définition suivante :

### Définition 20.111.

Un champ de vecteurs  $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  est un champ **conservatif** si il existe une fonction  $V: \mathbb{R}^3 \rightarrow \mathbb{R}$  telle que

$$F(x) = \nabla V(x). \quad (20.412)$$

Nous disons aussi parfois que le champ  $F$  dérive d'un potentiel ou bien qu'il s'agit d'un champ de gradient.

Les champs de vecteurs conservatifs sont particulièrement importants parce que presque toutes les forces connues en physiques dérivent d'un potentiel. Nous verrons que la terminologie « conservatif » provient du fait que les forces de ce type conservent l'énergie associée.

### Proposition 20.112.

Considérons une fonction  $V: \mathbb{R}^3 \rightarrow \mathbb{R}$  (que nous appellerons potentiel) et le champ de vecteur qui en dérive :

$$F = \nabla V. \quad (20.413)$$

Alors

$$\int_{\sigma} F \cdot ds = V(\sigma(b)) - V(\sigma(a)). \quad (20.414)$$

Autrement dit, le travail nécessaire pour déplacer un objet d'un point à un autre dans un champ de force conservatif vaut la différence de potentiel entre le point de départ et le point d'arrivée.

*Démonstration.* Par définition,

$$\int_{\sigma} F \cdot ds = \int_a^b F(\sigma(t)) \cdot \sigma'(t) dt. \quad (20.415)$$

Nous pouvons transformer l'intégrande de la façon suivante :

$$\begin{aligned} F(\sigma(t)) \cdot \sigma'(t) &= \nabla V(\sigma(t)) \cdot \sigma'(t) \\ &= \frac{\partial V}{\partial x}(\sigma(t))\sigma'_x(t) + \frac{\partial V}{\partial y}(\sigma(t))\sigma'_y(t) + \frac{\partial V}{\partial z}(\sigma(t))\sigma'_z(t) \\ &= \frac{d}{dt}[V(\sigma(t))] \end{aligned} \quad (20.416)$$

où nous avons posé

$$\sigma(t) = \begin{pmatrix} \sigma_x(t) \\ \sigma_y(t) \\ \sigma_z(t) \end{pmatrix} \quad (20.417)$$

et utilisé à l'envers la formule de dérivation de fonction composée pour

$$\frac{d}{dt}[V(\sigma(t))] = ((V \circ \sigma)(t))'. \quad (20.418)$$

En remettant ces expressions dans l'intégrale (20.415),

$$\int_{\sigma} F \cdot ds = \int_a^b \frac{d}{dt}[V(\sigma(t))] dt = V(\sigma(b)) - V(\sigma(a)). \quad (20.419)$$

□

### Exemple 20.113.

Nous savons que le champ de gravitation dérive d'un potentiel. À la surface de la Terre, le potentiel de gravitation vu par une masse  $m$  est donné par la fonction  $V(x, y, z) = mgz$ . Si nous voulons soulever cette masse d'une hauteur  $h$ , cela demandera toujours une énergie  $mgh$ , quel que soit le chemin suivi : en ligne droite vertical, en diagonal, en hélice, ...  $\triangle$

### Exemple 20.114.

À plus grande échelle, le champ de gravitation est encore un champ qui dérive d'un potentiel. En coordonnées sphériques,

$$V(\rho, \theta, \varphi) = k \frac{m}{\rho} \quad (20.420)$$

Lorsqu'un satellite a une orbite de rayon  $R$  autour la Terre, il reste sur la sphère  $\rho = R$ . Donc il reste sur une surface sur laquelle  $V$  est constante. Il n'y a donc pas de travail de la force de gravitation ! C'est pour cela qu'un satellite peut tourner pendant des siècles sans apport énergétique.  $\triangle$

### Exemple 20.115.

Soit le champ de vecteurs

$$F \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix} \quad (20.421)$$

et le chemin

$$\sigma(t) = \begin{pmatrix} t^4/4 \\ \sin^3(t\frac{\pi}{2}) \end{pmatrix} \quad (20.422)$$

Nous voulons calculer la circulation de  $F$  le long du chemin  $\sigma$  entre  $t = 0$  et  $t = 1$ .

La première chose à voir est que  $F = \nabla V$  avec  $V(x, y) = xy$ . Donc la circulation sera donnée par

$$\int_{\sigma} F \cdot ds = V(\sigma(1)) - V(\sigma(0)) = V\left(\frac{1}{4}, 1\right) - V(0, 0) = \frac{1}{4} - 0 = \frac{1}{4}. \quad (20.423)$$

Nous n'avons pas réellement calculé l'intégrale.  $\triangle$



## 20.19 Intégration de fonction à deux variables

### 20.19.1 Intégration sur un domaine rectangulaire

PgRapIntMultFubiniRect

Soit une fonction positive

$$\begin{aligned} f &: [a, b] \times [c, d] \rightarrow \mathbb{R}^+ \\ (x, y) &\mapsto f(x, y). \end{aligned} \quad (20.424)$$

L'intégrale de  $f$  sur le rectangle  $[a, b] \times [c, d]$  est le volume sous le graphe de la fonction. C'est-à-dire le volume de l'ensemble

$$\{(x, y, z) \text{ tel que } (x, y) \in [a, b] \times [c, d], z \leq f(x, y)\}. \quad (20.425)$$

**Théorème 20.116** (Théorème de Fubini).

Soit une fonction  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction continue par morceaux sur  $\mathcal{R} = [a, b] \times [c, d]$ . Alors

$$\int_{\mathcal{R}} f(x, y) dx dy = \int_a^b \left[ \int_c^d f(x, y) dy \right] dx = \int_c^d \left[ \int_a^b f(x, y) dx \right] dy. \quad (20.426)$$

En pratique, nous utilisons le théorème de Fubini pour calculer les intégrales sur des rectangles.

**Exemple 20.117.**

Nous voudrions intégrer la fonction  $f(x, y) = 4 + x^2 + y^2$  sur le rectangle de la figure 20.6.

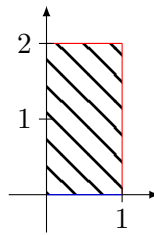


FIGURE 20.6: Intégration sur un rectangle

LabelFigVNBGooSqMsGU

L'ensemble sur lequel nous intégrons est donné par le produit cartésien d'intervalles  $E = [0, 1] \times [0, 2]$ . Le théorème de Fubini montre que nous pouvons intégrer séparément sur l'intervalle horizontal et vertical :

$$\int_{E=[0,1] \times [0,2]} f = \int_{[0,1]} \left( \int_{[0,2]} (4 - x^2 - y^2) dy \right) dx. \quad (20.427)$$

Ces intégrales sont maintenant des intégrales usuelles qui s'effectuent en calculant des primitives :

$$\begin{aligned} \int_0^1 \int_0^2 (4 - x^2 - y^2) dy dx &= \int_0^1 \left[ 4y - x^2 y - \frac{y^3}{3} \right]_0^2 dx \\ &= \int_0^1 \left( 8 - 2x^2 - \frac{8}{3} \right) dx \\ &= \left[ \frac{16x}{3} - \frac{2x^3}{3} \right]_0^1 \\ &= \frac{14}{3}. \end{aligned} \quad (20.428)$$

Avec Sage, on peut faire comme ceci :

```
-----
| Sage Version 4.6.1, Release Date: 2011-01-11                |
| Type notebook() for the GUI, and license() for information. |
-----
```

```
sage: f(x,y)=4-x**2-y**2
sage: f.integrate(y,0,2).integrate(x,0,1)
(x, y) |--> 14/3
```

△

### 20.19.2 Intégration sur un domaine non rectangulaire

Nous voulons maintenant intégrer la fonction  $f(x, y) = x^2 + y^2$  sur le triangle de la figure 20.7.

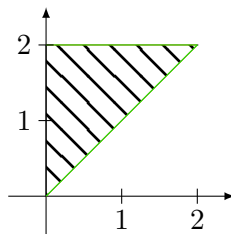
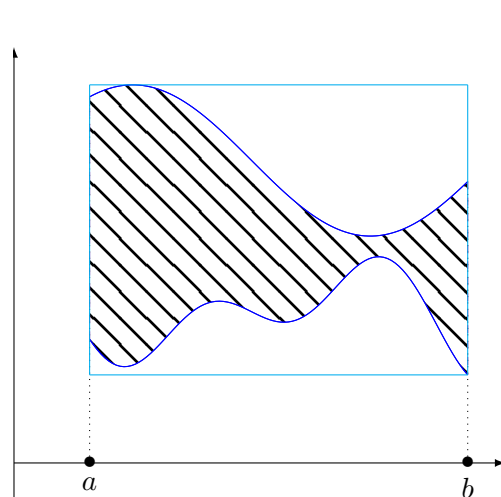


FIGURE 20.7: Intégration sur un triangle

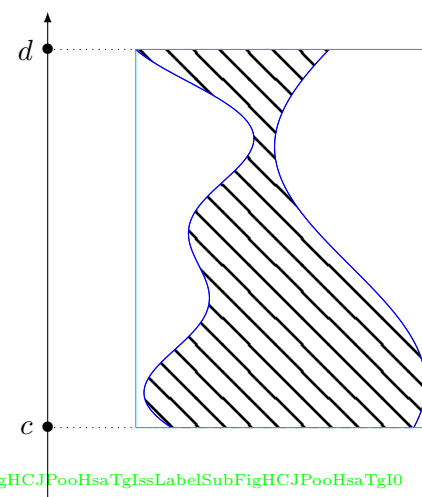
Étant donné que  $y$  varie de 0 à 2 et que *pour chaque*  $y$ , la variable  $x$  varie de 0 à  $y$ , nous écrivons l'intégrale sur le triangle sous la forme :

$$\int_{\text{triangle}} (x^2 + y^2) dx dy = \int_0^2 \left( \int_0^y (x^2 + y^2) dx \right) dy. \quad (20.429)$$

Il existe principalement deux types de domaines non rectangulaires : les « horizontaux » et les « verticaux », voir figure 20.8.



(a) Un domaine horizontal.



(b) Un domaine vertical.

FIGURE 20.8: Deux types de surfaces. Nous avons tracé un rectangle qui contient chacune des deux surfaces. L'intégrale sur un domaine sera l'intégrale sur le rectangle de la fonction qui vaut zéro en dehors du domaine.

Les surfaces horizontales sont de la forme

$$D = \{(x, y) \text{ tel que } x \in [a, b], \varphi_1(x) \leq y \leq \varphi_2(x)\} \quad (20.430)$$

où  $\varphi_1$  et  $\varphi_2$  sont les deux fonctions qui bornent le domaine. Le domaine  $D$  est la région comprise entre les graphes de  $\varphi_1$  et  $\varphi_2$ . Pour un tel domaine nous avons

$$\int_D f(x, y) dx dy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy. \quad (20.431)$$

Les surfaces verticales sont de la forme

$$D = \{(x, y) \text{ tel que } y \in [c, d], \psi_1(y) \leq x \leq \psi_2(y)\} \quad (20.432)$$

où  $\varphi_1$  et  $\varphi_2$  sont les deux fonctions qui bornent le domaine. Le domaine  $D$  est la région comprise entre les graphes de  $\varphi_1$  et  $\varphi_2$ . Dans ces cas nous avons

$$\int_D f = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx. \quad (20.433)$$

**Proposition 20.118.**

L'aire du domaine  $D$  vaut l'intégrale de la fonction  $f(x, y) = 1$  sur  $D$  :

$$\text{Aire}(D) = \int_D dx dy. \quad (20.434)$$

*Démonstration.* Supposons que le domaine soit du type « horizontal ». En utilisant le théorème de Fubini avec  $f(x, y) = 1$  nous avons

$$\int_D dx dy = \int_a^b \left[ \int_{\varphi_1(x)}^{\varphi_2(x)} dy \right] dx = \int_a^b [\varphi_2(x) - \varphi_1(x)] dx. \quad (20.435)$$

Cela représente l'aire sous  $\varphi_2$  moins l'aire sous  $\varphi_1$ , et par conséquent l'aire contenue entre les deux.  $\square$

**Exemple 20.119.**

Cherchons la surface du disque de centre  $(0, 0)$  et de rayon 1 dessinée à la figure 20.9.

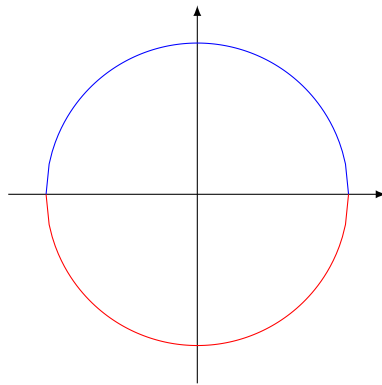


FIGURE 20.9: En bleu, la fonction  $\sqrt{r^2 - x^2}$  et en rouge, la fonction  $-\sqrt{r^2 - x^2}$ .

Le domaine est donné par  $\varphi_1(x) \leq y \leq \varphi_2(x)$  et  $x \in [-r, r]$  où  $\varphi_1(x) = -\sqrt{r^2 - x^2}$  et  $\varphi_2(x) = \sqrt{r^2 - x^2}$ . L'aire est donc donnée par

$$A = \int_{-r}^r [\varphi_2(x) - \varphi_1(x)] dx = 2 \int_{-r}^r \sqrt{r^2 - x^2} dx = 4 \int_0^r \sqrt{r^2 - x^2} dx. \quad (20.436)$$

Nous effectuons le premier changement de variables  $x = ru$ , donc  $dx = r du$ . En ce qui concerne les bornes, si  $x = 0$ , alors  $u = 0$  et si  $x = r$ , alors  $u = 1$ . L'intégrale à calculer devient

$$A = 4 \int_0^1 \sqrt{r^2 - r^2 u^2} r du = 4r^2 \int_0^1 \sqrt{1 - u^2} du. \quad (20.437)$$

Cette dernière intégrale se calcule en posant

$$\begin{aligned} u &= \sin(t) & du &= \cos(t)dt \\ u &= 0 & t &= 0 \\ u &= 1 & t &= \pi/2. \end{aligned} \quad (20.438)$$

Nous avons

$$A = 4r^2 \int_0^{\pi/2} \sqrt{1 - \sin^2(t)} \cos(t) dt = 4r^2 \int_0^{\pi/2} \cos^2(t) dt. \quad (20.439)$$

En utilisant la formule  $2 \cos^2(x) = 1 + \cos(2x)$ , nous avons

$$A = 4r^2 \int_0^{\pi/2} \frac{1 + \cos(2t)}{2} dt = \pi r^2. \quad (20.440)$$

△

### 20.19.3 Changement de variables

Nous n'allons pas parler de changements de variables maintenant parce que les principaux exemples sont les coordonnées polaires, cylindriques et sphériques qui requièrent les fonctions trigonométriques. Ce sera pour la section 18.15.

## 20.20 Les intégrales triples

Les intégrales triples fonctionnent exactement de la même manière que les intégrales doubles. Il s'agit de déterminer sur quelle domaine les variables varient et d'intégrer successivement par rapport à  $x$ ,  $y$  et  $z$ . Il est autorisé de permuter l'ordre d'intégration<sup>26</sup> à condition d'adapter les domaines d'intégration.

### Exemple 20.120.

Soit le domaine parallélépipédique rectangle

$$R = [0, 1] \times [1, 2] \times [0, 4]. \quad (20.441)$$

Pour intégrer la fonction  $f(x, y, z) = x^2 y \sin(z)$  sur  $R$ , nous faisons

$$\begin{aligned} I &= \int_R x^2 y \sin(z) dx dy dz \\ &= \int_0^1 dx \int_1^2 dy \int_0^4 x^2 y \sin(z) dz \\ &= \int_0^1 dx \int_1^2 x^2 y (1 - \cos(4)) dy \\ &= \int_0^1 \frac{3}{2} (1 - \cos(4)) x^2 dx \\ &= \frac{1}{2} (1 - \cos(4)). \end{aligned} \quad (20.442)$$

---

```
| Sage Version 4.6.1, Release Date: 2011-01-11 |
| Type notebook() for the GUI, and license() for information. |
|-----|
```

```
sage: f(x,y,z)=x**2*y*sin(z)
sage: f.integrate(x,0,1).integrate(y,1,2).integrate(z,0,4)
(x, y, z) |--> -1/2*cos(4) + 1/2
```

26. En toute rigueur, cela n'est pas vrai, mais nous ne considérons seulement des cas où cela est autorisé.

**Exemple 20.121.**

Soit  $D$  la région délimitée par le plan  $x = 0$ ,  $y = 0$ ,  $z = 2$  et la surface d'équation

$$z = x^2 + y^2. \quad (20.443)$$

Cherchons à calculer  $\int_D x \, dx \, dy \, dz$ . Ici, un dessin indique que le volume considéré est  $z \geq x^2 + y^2$ . Il y a plusieurs façons de décrire cet ensemble. Une est celle-ci :

$$\begin{aligned} z &: 0 \rightarrow 2 \\ x &: 0 \rightarrow \sqrt{z} \\ y &: 0 \rightarrow \sqrt{z - x^2}. \end{aligned} \quad (20.444)$$

Cela revient à dire que  $z$  peut prendre toutes les valeurs de 0 à 2, puis que pour chaque  $z$ , la variable  $x$  peut aller de 0 à  $\sqrt{z}$ , mais que pour chaque  $z$  et  $x$  fixés, la variable  $y$  ne peut pas dépasser  $\sqrt{z - x^2}$ . En suivant cette méthode, l'intégrale à calculer est

$$\int_0^2 dz \int_0^{\sqrt{z}} dx \int_0^{\sqrt{z-x^2}} f(x, y, z) dy. \quad (20.445)$$

```
-----
| Sage Version 4.6.1, Release Date: 2011-01-11 |
| Type notebook() for the GUI, and license() for information. |
-----
sage: f(x,y,z)=x
sage: assume(z>0)
sage: assume(z-x**2>0)
sage: f.integrate(y,0,sqrt(z-x**2)).integrate(x,0,sqrt(z)).integrate(z,0,2)
(x, y, z) |--> 8/15*sqrt(2)
```

Notez qu'il a fallu aider Sage en lui indiquant que  $z > 0$  et  $z - x^2 > 0$ .

Un autre paramétrage serait

$$\begin{aligned} x &: 0 \rightarrow \sqrt{2} \\ y &: 0 \rightarrow \sqrt{2 - x^2} \\ z &: x^2 + y^2 \rightarrow 2. \end{aligned} \quad (20.446)$$

```
-----
| Sage Version 4.6.1, Release Date: 2011-01-11 |
| Type notebook() for the GUI, and license() for information. |
-----
sage: f(x,y,z)=x
sage: assume(2-x**2>0)
sage: f.integrate(y,0,sqrt(2-x**2)).integrate(x,0,sqrt(2)).integrate(z,0,2)
(x, y, z) |--> 8/15*sqrt(2)
```

Écrivons le détail de cette dernière intégrale :

$$\begin{aligned} I &= \int_0^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} dy \int_{x^2+y^2}^2 x \, dz \\ &= \int_0^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} x(2 - x^2 - y^2) dy \\ &= \int_0^{\sqrt{2}} dx x \left[ (2 - x^2)y - \frac{y^3}{3} \right]_0^{\sqrt{2-x^2}} \\ &= \int_0^{\sqrt{2}} \frac{2}{3} x(2 - x^2)^{3/2} dx. \end{aligned} \quad (20.447)$$

Ici nous effectuons le changement de variable  $u = x^2$ ,  $du = 2xdx$ . Ne pas oublier de changer les bornes de l'intégrale :

$$I = \frac{1}{3} \int_0^2 (2-u)^{3/2} du. \quad (20.448)$$

Le changement de variable  $t = 2 - u$ ,  $dt = -du$  fait venir (attention aux bornes!!)

$$I = -\frac{1}{3} \int_2^0 t^{3/2} dt = \frac{1}{3} \left[ \frac{t^{5/2}}{5/2} \right]_0^2 = \frac{8}{15} \sqrt{2}. \quad (20.449)$$

△

### 20.20.1 Volume

Parmi les nombreuses interprétations géométriques de l'intégrale triple, notons celle-ci :

**Proposition 20.122.**

Soit  $D \subset \mathbb{R}^3$ . Le volume de  $D$  est donné par

$$\text{Vol}(D) = \int_D dx dy dz. \quad (20.450)$$

C'est-à-dire l'intégrale de la fonction  $f(x, y, z) = 1$  sur  $D$ .

Suivant les points de vue, cette proposition peut être considérée comme une *définition* du volume.

ExemVolSphCart

**Exemple 20.123.**

Calculons le volume de la sphère de rayon  $R$ . Le domaine de variation des variables  $x$ ,  $y$  et  $z$  pour la sphère est

$$\begin{aligned} x: & -R \rightarrow R \\ y: & -\sqrt{R^2 - x^2} \rightarrow \sqrt{R^2 - x^2} \\ z: & -\sqrt{R^2 - x^2 - y^2} \rightarrow \sqrt{R^2 - x^2 - y^2}. \end{aligned} \quad (20.451)$$

Par conséquent nous devons calculer l'intégrale

$$V = \int_{-R}^R dx \int_{-\sqrt{R^2 - x^2}}^{\sqrt{R^2 - x^2}} dy \int_{-\sqrt{R^2 - x^2 - y^2}}^{\sqrt{R^2 - x^2 - y^2}} dz. \quad (20.452)$$

La première intégrale est simple :

$$V = 2 \int_{-R}^R dx \int_{-\sqrt{R^2 - x^2}}^{\sqrt{R^2 - x^2}} \sqrt{R^2 - x^2 - y^2} dy. \quad (20.453)$$

Afin de simplifier la notation, nous posons  $a = R^2 - x^2$ . Ceci n'est pas un changement de variables : juste une notation provisoire le temps d'effectuer l'intégration sur  $y$ . Étudions donc

$$I = \int_{-\sqrt{a}}^{\sqrt{a}} \sqrt{a - y^2} dy, \quad (20.454)$$

ce qui est la surface du demi-disque de rayon  $\sqrt{a}$ . Nous avons donc

$$I = \frac{\pi a}{2} = \frac{\pi}{2} (R^2 - x^2), \quad (20.455)$$

et

$$V = 2 \int_{-R}^R \frac{\pi}{2} (R^2 - x^2) dx = \pi \left[ R^2 x - \frac{x^3}{3} \right]_{-R}^R = \frac{4}{3} \pi R^3. \quad (20.456)$$

△

**Exemple 20.124.**

Nous pouvons calculer le volume de la sphère en utilisant les coordonnées sphériques. Les bornes des variables pour la sphère de rayon  $R$  sont

$$\begin{aligned}\rho &: 0 \rightarrow R \\ \theta &: 0 \rightarrow \pi \\ \varphi &: 0 \rightarrow 2\pi.\end{aligned}\tag{20.457}$$

En n'oubliant pas le jacobien  $\rho^2 \sin(\theta)$ , l'intégrale à calculer est

$$V = \int_0^R d\rho \int_0^{2\pi} d\varphi \int_0^\pi \rho^2 \sin(\theta) d\theta\tag{20.458}$$

L'intégrale sur  $\varphi$  fait juste une multiplication par  $2\pi$ . Celle sur  $\rho$  vaut

$$\int_0^R \rho^2 d\rho = \frac{R^3}{3}.\tag{20.459}$$

L'intégrale sur  $\theta$  donne

$$\int_0^\pi \sin(\theta) d\theta = [-\cos(\theta)]_0^\pi = 2.\tag{20.460}$$

Le tout fait par conséquent

$$V = \frac{4}{3}\pi R^3.\tag{20.461}$$

Sans contestes, le passage aux coordonnées sphériques a considérablement simplifié le calcul par rapport à celui de l'exemple 20.123.  $\triangle$

## 20.21 Un petit peu plus formel

### 20.21.1 Intégration sur un domaine non rectangulaire

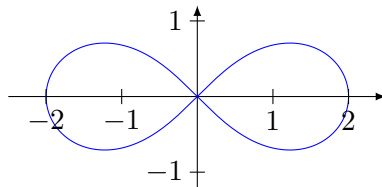


FIGURE 20.10: Intégrer sur des domaines plus complexes.

LabelFigPONXooXYjEot

La méthode de Fubini ne fonctionne plus sur un domaine non rectangulaire tel que celui de la figure 20.10. Nous allons donc utiliser une astuce. Considérons le domaine

$$E = \{(x, y) \in \mathbb{R}^2 \text{ tel que } a < x < b \text{ et } \alpha(x) < y < \beta(x)\}\tag{20.462}$$

représenté sur la figure 20.10. Nous considérons la fonction

$$\tilde{f}(x, y) = \begin{cases} f(x, y) & \text{si } (x, y) \in E \\ 0 & \text{sinon.} \end{cases}\tag{20.463}$$

Ensuite intégrons  $\tilde{f}$  sur un rectangle qui englobe la surface à intégrer à l'aide de Fubini. Étant donné que  $\tilde{f} = f$  sur la surface et que  $\tilde{f}$  est nulle en dehors, nous avons

$$\int_E f = \int_E \tilde{f} = \int_{\text{rectangle}} \tilde{f} = \int_a^b \left( \int_{\alpha(x)}^{\beta(x)} f(x, y) dy \right) dx.\tag{20.464}$$

Dans le cas de l'intégrale de  $f(x, y) = x^2 + y^2$  sur le triangle de la figure 20.7, nous avons

$$\int_{\text{triangle}} (x^2 + y^2) dx dy = \int_0^2 \left( \int_0^y (x^2 + y^2) dx \right) dy.\tag{20.465}$$

**Remarque 20.125.**

Le nombre  $\int_D f(x, y) dx dy$  ne dépend pas du choix du rectangle englobant  $D$ .

En pratique, nous calculons l'intégrale en utilisant une extension du théorème de Fubini :

**Théorème 20.126.**

Soit  $f: D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction continue où  $D$  est un domaine de type vertical ou horizontal.

(1) Si  $D$  est vertical, alors

$$\int_D f = \int_a^b \left[ \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right] dx. \quad (20.466)$$

(2) Si  $D$  est horizontal, alors

$$\int_D f = \int_c^d \left[ \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) dx \right] dy. \quad (20.467)$$

**20.21.1.1 Coordonnées polaires**

Les coordonnées polaires sont données par le difféomorphisme

$$\begin{aligned} g: ]0, \infty[ \times ]0, 2\pi[ &\rightarrow \mathbb{R}^2 \setminus D \\ (r, \theta) &\mapsto (r \cos(\theta), r \sin(\theta)) \end{aligned} \quad (20.468)$$

où  $D$  est la demi-droite  $y = 0, x \geq 0$ . Le fait que les coordonnées polaires ne soient pas un difféomorphisme sur tout  $\mathbb{R}^2$  n'est pas un problème pour l'intégration parce que le manque de difféomorphisme est de mesure nulle dans  $\mathbb{R}^2$ . Le jacobien est donné par

$$Jg = \det \begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix} = r. \quad (20.469)$$

**20.21.1.2 Coordonnées sphériques**

Les coordonnées sphériques sont données par

$$\begin{cases} x = r \cos \theta \sin \varphi \\ y = r \sin \theta \sin \varphi \\ z = r \cos \varphi \end{cases} \quad \text{avec} \quad \begin{aligned} r &\in ]0, \infty[ \\ \theta &\in ]0, 2\pi[ \\ \varphi &\in ]0, \pi[. \end{aligned} \quad (20.470)$$

Le jacobien associé est  $Jg(r, \theta, \varphi) = -r^2 \sin \varphi$ . Rappelons que ce qui rentre dans l'intégrale est la valeur absolue du jacobien.

Si nous voulons calculer le volume de la sphère de rayon  $R$ , nous écrivons donc

$$\int_0^R dr \int_0^{2\pi} d\theta \int_0^\pi r^2 \sin(\varphi) d\varphi = 4\pi R = \frac{4}{3}\pi R^3. \quad (20.471)$$

Ici, la valeur absolue n'est pas importante parce que lorsque  $\varphi \in ]0, \pi[,$  le sinus de  $\varphi$  est positif.

Des petits malins pourraient remarquer que le changement de variable (20.470) est encore un paramétrage de  $\mathbb{R}^3$  si on intervertit le domaine des angles :

$$\begin{aligned} \theta &: 0 \rightarrow \pi \\ \varphi &: 0 \rightarrow 2\pi, \end{aligned} \quad (20.472)$$

alors nous paramétrons encore parfaitement bien la sphère, mais hélas

$$\int_0^R dr \int_0^\pi d\theta \int_0^{2\pi} r^2 \sin(\varphi) d\varphi = 0. \quad (20.473)$$



Pourquoi ces « nouvelles » coordonnées sphériques sont-elles mauvaises ? Il y a que quand l'angle  $\phi$  parcourt  $]0, 2\pi[$ , son sinus n'est plus toujours positif, donc la *valeur absolue* du jacobien n'est plus  $r^2 \sin(\phi)$ , mais  $r^2 \sin(\phi)$  pour les  $\phi$  entre 0 et  $\pi$ , puis  $-r^2 \sin(\phi)$  pour  $\phi$  entre  $\pi$  et  $2\pi$ . Donc l'intégrale (20.473) n'est pas correcte. Il faut la remplacer par

$$\int_0^R dr \int_0^\pi d\theta \int_0^\pi r^2 \sin(\phi) d\phi - \int_0^R dr \int_0^\pi d\theta \int_\pi^{2\pi} r^2 \sin(\phi) d\phi = \frac{4}{3}\pi R^3 \quad (20.474)$$

## 20.22 Aire et primitive

Soient  $f: \mathbb{R} \rightarrow \mathbb{R}$  une fonction continue et  $x \in \mathbb{R}$ . Pour chaque  $a \in \mathbb{R}$ , nous pouvons considérer le nombre  $F(x)$  défini par

$$F(x) = \int_a^x f(t) dt. \quad (20.475)$$

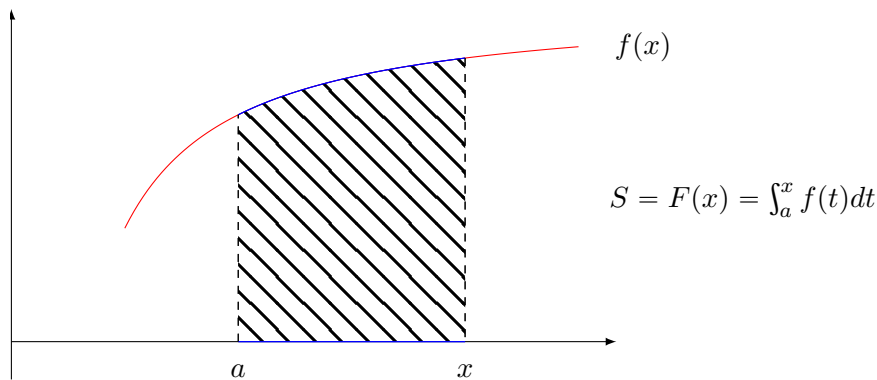


FIGURE 20.11: Surface sous une courbe

LabelFigVSZRooRWgUGu

La fonction  $F$  ainsi définie a deux importantes propriétés :

- (1) C'est une primitive de  $f$ ,
- (2) Elle donne la surface en dessous de  $f$  entre les points  $a$  et  $x$ , voir la figure 20.11.

Notons que tant que  $f$  est positive, la surface est croissante.

La manière de calculer la surface comprise entre deux fonctions est dessinée à la figure 20.12.

La surface entre les deux fonctions  $y_1(x)$  et  $y_2(x)$  se calcule comme suit.

- (1) On calcule les intersections entre  $y_1$  et  $y_2$ . Notons  $a$  et  $b$  les ordonnées obtenues.
- (2) La surface demandée est la différence entre la surface sous la fonction  $y_1$  (la plus grande) et la surface sous la fonction  $y_2$  (la plus petite), donc

$$S = \int_a^b y_1 - \int_a^b y_2. \quad (20.476)$$

### 20.22.1 Longueur d'arc de courbe

La longueur de l'arc de courbe de la fonction  $y = f(x)$  entre les abscisses  $x_0$  et  $x_1$  est donné par la formule

$$l(x_0, x_1) = \int_{x_0}^{x_1} \sqrt{1 + y'(t)^2} dt. \quad (20.477) \quad \text{EqLongArcCourbe}$$

Lorsque la courbe est donnée sous forme paramétrique

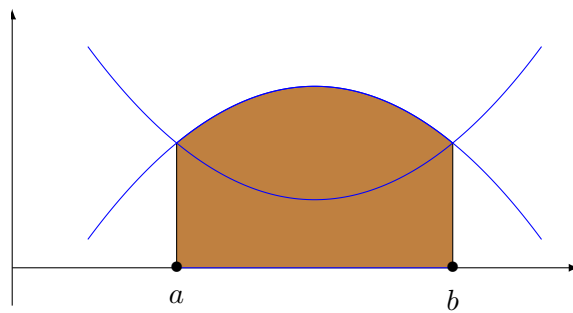
$$\begin{cases} x = x(t) \\ y = y(t), \end{cases} \quad (20.478a)$$

$$(20.478b)$$

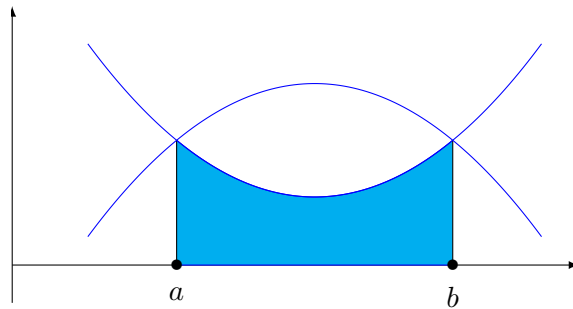
alors la formule devient

$$l(t_1, t_2) = \int_{t_1}^{t_2} \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt, \quad (20.479) \quad \text{EqLongArcParam}$$

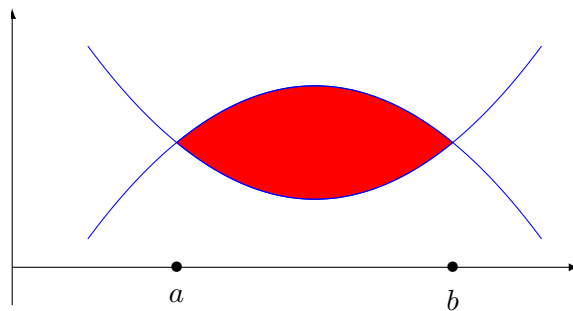
où  $\dot{x}(t) = x'(t)$ .



(a) La grande surface.



(b) La surface à soustraire.



(c) La surface entre les deux fonctions.

FIGURE 20.12: Le calcul de la surface comprise entre deux fonctions.

### 20.22.2 Aire de révolution

Pour savoir l'aire engendrée par la ligne  $y = f(x)$  entre  $a$  et  $b$  autour de l'axe  $Ox$ , on utilise la formule

$$S = 2\pi \int_a^b \sqrt{1 + f'(x)^2} f(x) dx. \quad (20.480)$$

## 20.23 L'aire en dessous d'une courbe

Soit  $f$  une fonction à valeurs dans  $\mathbb{R}^+$ .

Nous voudrions pouvoir calculer l'aire au-dessous du graphe de la fonction  $f$ . Nous notons  $S_f(x)$  l'aire là-dessous de la fonction  $f$  entre l'abscisse 0 et  $x$ , c'est-à-dire l'aire bleue de la figure 20.13.

Si la fonction  $f$  est continue et que  $\Delta x$  est assez petit, la fonction ne varie pas beaucoup entre  $x$  et  $x + \Delta x$ . L'augmentation de surface entre  $x$  et  $x + \Delta x$  peut donc être approximée par le rectangle de surface  $f(x)\Delta x$ . Ce que nous avons donc, c'est que quand  $\Delta x$  est très petit,

$$S_f(x + \Delta x) - S_f(x) = f(x)\Delta x, \quad (20.481)$$

ou encore

$$f(x) = \frac{S_f(x + \Delta x) - S_f(x)}{\Delta x}. \quad (20.482)$$

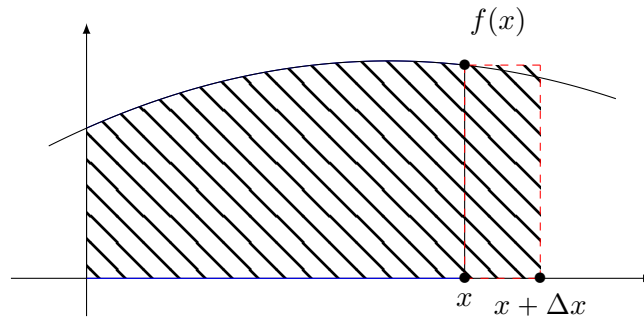


FIGURE 20.13: L'aire en dessous d'une courbe. Le rectangle rouge d'aire  $f(x)\Delta x$  approxime de combien la surface augmente lorsqu'on passe de  $x$  à  $x + \Delta x$ .

LabelFigKKRooHseDzC

Nous formalisons la notion de « lorsque  $\Delta x$  est très petit » par une limite :

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{S_f(x + \Delta x) - S_f(x)}{\Delta x}. \quad (20.483)$$

Donc, la fonction  $f$  est la dérivée de la fonction qui représente l'aire là-dessous de  $f$ . Calculer des surfaces revient donc au travail inverse de calculer des dérivées.

## 20.24 Propriétés des intégrales

LemIneqnormeintintnorm

### Lemme 20.127.

Pour toute fonction  $u: [a, b] \rightarrow \mathbb{R}^n$ , nous avons

$$\left\| \int_a^b u(t) dt \right\| \leq \int_a^b \|u(t)\| dt \quad (20.484)$$

pourvu que le membre de gauche ait un sens.

*Démonstration.* Étant donné que  $\int_a^b u(t) dt$  est un élément de  $\mathbb{R}^n$ , par la proposition 10.19, il existe un  $\xi \in \mathbb{R}^n$  de norme 1 tel que

$$\left\| \int_a^b u(t) dt \right\| = \xi \cdot \int_a^b u(t) dt = \int_a^b u(t) \cdot \xi dt \leq \int_a^b \|u(t)\| \|\xi\| dt = \int_a^b \|u(t)\| dt. \quad (20.485)$$

□

### Proposition 20.128 (Relations de Chasles).

Soit  $f$  une fonction continue sur l'intervalle  $I$ . Si  $a, b, c \in I$  nous avons

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx. \quad (20.486)$$

Sur la figure 20.14, la surface de  $a$  à  $c$  est évidemment égale à la somme des surfaces de  $a$  à  $b$  et de  $b$  à  $c$ .

### Corolaire 20.129.

$$\int_a^b f(x) dx = - \int_b^a f(x) dx. \quad (20.487)$$

lineariteintegrale

### Proposition 20.130 (Linéarité de l'intégrale).

Si  $f$  et  $g$  sont deux fonctions continues sur  $I \subset \mathbb{R}$ ,  $a, b \in I$  et  $\lambda \in \mathbb{R}$  nous avons

$$\int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x) dx, \quad (20.488)$$

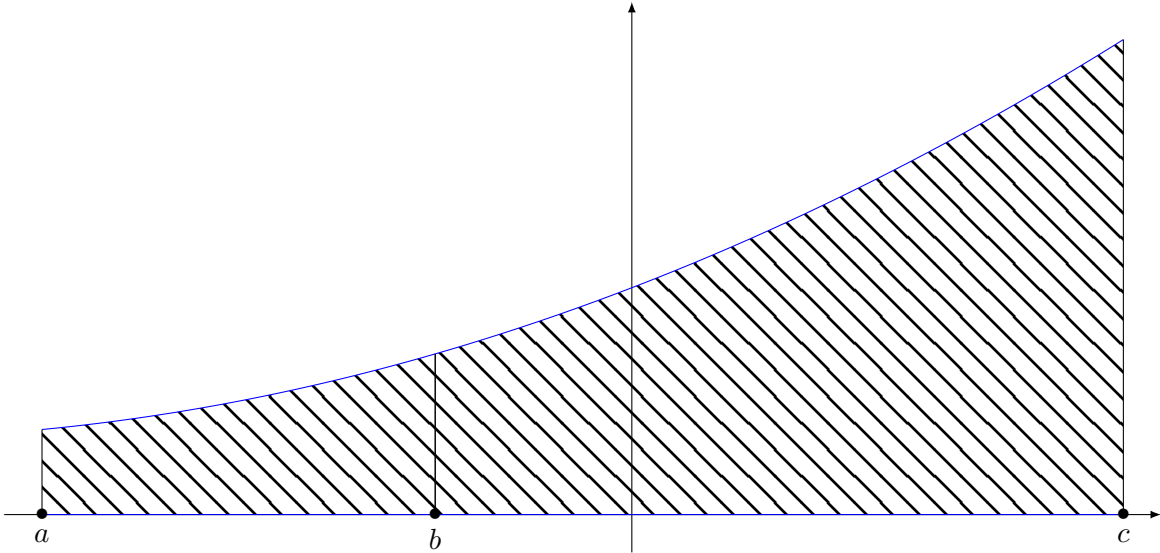


FIGURE 20.14: Illustration pour les relations de Chasles. LabelFigNWDooOObSHB

et

$$\int_a^b \lambda f(x) dx = \lambda \int_a^b f(x) dx. \tag{20.489}$$

PropCJIooHqECbq

**Proposition 20.131** (L'intégrale est monotone).  
Soient  $a, b \in I$  avec  $a < b$ . Si  $f \geq g$  sur  $[a, b]$  alors

$$\int_a^b f(x) dx \geq \int_a^b g(x) dx. \tag{20.490}$$

PropHVWooBDRhCX

**Corolaire 20.132** (Positivité).  
Si  $a < b$  et  $f \geq 0$  sur  $[a, b]$  alors

$$\int_a^b f(x) dx \geq 0. \tag{20.491}$$

Ce résultat n'est qu'une application de la proposition 20.131 car il consiste à prendre comme fonction  $g$  la fonction nulle.

### 20.25 Techniques d'intégration

Par le théorème 14.288, le calcul d'une intégrale consiste essentiellement à trouver une primitive de la fonction à intégrer. Il est donc indispensable de bien connaître les dérivées des fonctions usuelles.

Voici un tableau des primitives à connaître.

PageLCHooMbWjOj

| Fonction<br>$f(x)$        | Primitive<br>$\int f(x) dx$         | Ensemble de définition<br>de $f$                                   | Remarques                                |
|---------------------------|-------------------------------------|--------------------------------------------------------------------|------------------------------------------|
| $x^\alpha$                | $\frac{x^{\alpha+1}}{\alpha+1} + C$ | dépend de $\alpha$                                                 | $\alpha \in \mathbb{R} \setminus \{-1\}$ |
| $\frac{1}{x}$             | $\ln( x ) + C$                      | $x \neq 0$                                                         |                                          |
| $\frac{1}{1+x^2}$         | $\arctan(x) + C$                    | $\mathbb{R}$                                                       |                                          |
| $\frac{1}{\sqrt{1-x^2}}$  | $\arcsin(x) + C$                    | $] -1, 1[$                                                         |                                          |
| $\frac{-1}{\sqrt{1-x^2}}$ | $\arccos(x) + C$                    | $] -1, 1[$                                                         |                                          |
| $e^x$                     | $e^x + C$                           | $\mathbb{R}$                                                       |                                          |
| $\sin(x)$                 | $-\cos(x) + C$                      | $\mathbb{R}$                                                       |                                          |
| $\cos(x)$                 | $\sin(x) + C$                       | $\mathbb{R}$                                                       |                                          |
| $1 + \tan^2(x)$           | $\tan(x) + C$                       | un intervalle de la forme $]-\frac{\pi}{2}, \frac{\pi}{2}[ + k\pi$ |                                          |

Notez que au signe près, les fonctions arcsin et arccos ont la même dérivée.

Si la fonction à intégrer est une combinaison linéaire de fonctions usuelles alors sa primitive peut être calculée en utilisant la proposition 20.130. Dans les sections suivantes on abordera deux autres cas où la fonction à intégrer peut s'écrire en termes de fonctions dont on connaît une primitive.

### 20.25.1 Intégration par parties

PROPooRLFIooQHnyJY

**Proposition 20.133** ([1]).

Si  $u$  et  $v$  sont deux fonctions dérivables de dérivées continues sur l'intervalle  $[a, b]$  alors

$$\int_a^b u(x)v'(x)dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx. \quad (20.492)$$

EQooKISBooQvGMQT

Si de plus les fonctions  $uv'$  et  $u'v$  sont intégrables sur  $[a, \infty]$ , alors

$$\int_a^\infty u(x)v'(x)dx = [u(x)v(x)]_a^\infty - \int_a^\infty u'(x)v(x)dx. \quad (20.493)$$

*Démonstration.* Il s'agit d'utiliser à l'envers la formule de dérivation d'un produit :

$$uv' = (uv)' - u'v. \quad (20.494)$$

Les fonctions à gauche et à droite étant égales, elles ont même intégrale sur  $[a, b]$  et par linéarité, voir proposition 20.130, on a :

$$\int_a^b u(x)v'(x)dx = \int_a^b (uv)'(x) - \int_a^b u'(x)v(x)dx. \quad (20.495)$$

La fonction  $uv$  est évidemment une primitive de  $(uv)'$ , de telle sorte que l'on puisse un peu simplifier cette expression :

$$\int_a^b u(x)v'(x)dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx, \quad (20.496)$$

ce qu'il fallait démontrer.

En ce qui concerne l'affirmation avec  $b = \infty$ , le lemme 14.262 est applicable et il suffit de passer à la limite dans (20.492).  $\square$

ExWIEooVUgvSp

**Exemple 20.134.**

Un cas typique d'utilisation de l'intégrale par parties est le suivant. Soit à calculer

$$\int_0^\pi x \cos(x)dx. \quad (20.497)$$

Nous devons écrire  $x \cos(x)$  comme un produit  $u(x)v'(x)$ . Il y a (au moins) deux moyens de le faire :

$$\begin{cases} u = x \\ v' = \cos(x). \end{cases} \quad (20.498a)$$

$$\begin{cases} u = \cos(x) \\ v' = x. \end{cases} \quad (20.498b)$$

ou

$$\begin{cases} u = \cos(x) \\ v' = x. \end{cases} \quad (20.499a)$$

$$\begin{cases} u = x \\ v' = \cos(x). \end{cases} \quad (20.499b)$$

Nous allons choisir le premier<sup>27</sup>. Nous avons donc

$$\begin{aligned} u &= x, & v' &= \cos(x) \\ u' &= 1 & v &= \sin(x). \end{aligned} \quad (20.500)$$

En utilisant la formule d'intégration par parties,

$$\int_0^\pi x \cos(x)dx = [x \sin(x)]_0^\pi - \int_0^\pi 1 \times \sin(x)dx = \pi \sin(\pi) - [-\cos(x)]_0^\pi = -2. \quad (20.501)$$

$\triangle$

27. Mais nous conseillons vivement au lecteur d'essayer le deuxième pour se rendre compte qu'il ne fonctionne pas.

Le plus souvent, pour alléger les notations, il est plus pratique d'utiliser l'intégration par parties pour déterminer une primitive. Nous utilisons pour cela la formule (sans doute plus simple à retenir)

$$\int uv' = uv - \int u'v. \quad (20.502)$$

**Exemple 20.135.**

Nous reprenons l'exemple 20.134 en déterminant cette fois une primitive de  $x \cos(x)$  :

$$\int x \cos(x) dx = x \sin(x) - \int \sin(x) dx = x \sin(x) + \cos(x) + C, \quad C \in \mathbb{R}. \quad (20.503)$$

Nous retrouvons le résultat numérique de l'exemple précédent en ajoutant les extrêmes d'intégration

$$\int_0^\pi x \cos(x) dx = [x \sin(x) + \cos(x)]_0^\pi = -2. \quad (20.504)$$

△

**Remarque 20.136.**

Lorsqu'on calcule des intégrales, il est bon de passer par la primitive (c'est-à-dire en suivant l'exemple 20.135 et non 20.134) parce qu'il est alors facile de vérifier le résultat en calculant la dérivée de la primitive trouvée.

Par exemple pour vérifier si (20.503) est correct, il suffit de dériver  $x \sin(x) + \cos(x)$  :

$$(x \sin(x) + \cos(x))' = \sin(x) + x \cos(x) - \sin(x) = x \cos(x). \quad (20.505)$$

La fonction  $x \sin(x) + \cos(x)$  est donc bien une primitive de  $x \cos(x)$ .

**Exemple 20.137** (Primitive du logarithme).

La primitive de la fonction logarithme définie en 15.91 nous offre un bon moment d'intégration par partie.

Trouver la primitive de la fonction  $x \mapsto \ln(x)$ . Pour calculer

$$\int \ln(x) dx \quad (20.506)$$

nous écrivons  $\ln(x) = 1 \times \ln(x)$  et nous posons  $u' = 1$  et  $v = \ln(x)$ , c'est-à-dire

$$\begin{aligned} u' &= 1 & v &= \ln(x) \\ u &= x & v' &= \frac{1}{x}. \end{aligned} \quad (20.507)$$

La formule d'intégration par parties (20.492) donne donc

$$\int \ln(x) = x \ln(x) - \int x \times \frac{1}{x} = x \ln(x) - \int 1 = x \ln(x) - x + C, \quad C \in \mathbb{R}. \quad (20.508)$$

Il est facile de vérifier par un petit calcul que

$$(x \ln(x) - x)' = \ln(x). \quad (20.509)$$

△

## 20.25.2 Changement de variables – pour trouver des primitives

De la même manière que l'utilisation « à l'envers » de la formule de dérivation du produit avait donné la méthode d'intégration par parties, nous allons voir que l'utilisation « à l'envers » de la formule de dérivation d'une fonction composée donne lieu à la méthode d'intégration par changement de variables.

PROPooMVIUooZmvHxS

**Proposition 20.138.**

Soient  $I$  et  $J$  des intervalles de  $\mathbb{R}$ ,  $u: I \rightarrow J$  une fonction qui est dérivable de dérivée continue et  $f: J \rightarrow \mathbb{R}$  une fonction admettant une primitive  $F$ . Alors la fonction

$$x \mapsto F(u(x)) \quad (20.510)$$

est une primitive de

$$f(u(x))u'(x). \quad (20.511) \text{changvar}$$

*Démonstration.* Cela est une utilisation immédiate de la formule de dérivée des fonctions composées.  $\square$

**Exemple 20.139.**

Soit à calculer

$$\int x\sqrt{1-x^2}dx. \quad (20.512)$$

La fonction  $g(x) = x\sqrt{1-x^2}$  est le produit de  $x$  et de  $\sqrt{1-x^2}$ . On remarque que la dérivée de  $1-x^2$  est  $-2x$  : nous avons alors, à un facteur  $-2$  près, une expression de la forme (20.511) où la racine carrée joue le rôle de  $f$ ,  $f(t) = \sqrt{t}$ , et  $1-x^2$  le rôle de  $u$ . Une primitive de la fonction  $f(t) = \sqrt{t}$  est  $F(t) = 2t^{3/2}/3$ .

Donc la fonction  $\frac{2u(x)^{3/2}}{3} = \frac{2}{3}(1-x^2)^{3/2}$  est primitive de  $-2x\sqrt{1-x^2} = -2g(x)$ . Autrement dit,

$$\int -2x\sqrt{1-x^2}dx = \frac{2(1-x^2)^{3/2}}{3} + C, \quad (20.513)$$

et en divisant par  $-2$  nous trouvons la primitive demandée :

$$\int x\sqrt{1-x^2}dx = -\frac{(1-x^2)^{3/2}}{3} + C. \quad (20.514)$$

$\triangle$

L'exemple suivant donne une façon plus économe de retenir la méthode du changement de variables.

exempleprimitivechangvar

**Exemple 20.140.**

Soit à calculer

$$\int \cos(x)e^{\sin(x)}dx. \quad (20.515)$$

Vu qu'il y a beaucoup de fonctions trigonométriques dans la fonction à intégrer, nous allons poser  $u(x) = \sin(x)$ , et remplacer élément par élément tout ce qui contient du «  $x$  » dans l'intégrale demandée par la quantité correspondante en termes de  $u$ .

La difficulté est de savoir ce que nous allons faire du «  $dx$  » dans l'intégrale. Ce  $dx$  marque une variation (infinitésimale) de  $x$ . La formule des accroissements finis dit que si  $x$  augmente de la valeur  $dx$ , alors  $u(x)$  augmente de  $u'(x)dx$ , c'est-à-dire que

$$du = \cos(x)dx. \quad (20.516)$$

Nous avons donc les substitutions suivantes à faire :

$$\sin(x) = u \quad (20.517a)$$

$$du = \cos(x)dx \quad (20.517b)$$

$$dx = \frac{du}{\cos(x)}. \quad (20.517c)$$

La chose « magique » est que le  $\cos(x)$  se trouvant dans la fonction se simplifie avec le cosinus qui arrive lorsqu'on remplace  $dx$  par  $\frac{du}{\cos(x)}$ . Les substitutions faites nous restons avec

$$\int \cos(x)e^{\sin(x)}dx = \int e^u du = e^u + C, \quad \text{où } u = \sin(x). \quad (20.518)$$

Attention : la réponse doit être impérativement donnée en termes de  $x$  et non de  $u$ . Nous écrivons donc

$$\int \cos(x)e^{\sin(x)} = e^{\sin(x)} + C. \quad (20.519)$$

△

### 20.25.3 Changement de variables – pour calculer des intégrales

Le théorème 14.288 fixe la relation entre la recherche des primitives de  $f$  et la calcul de l'intégrale de  $f$  sur l'intervalle d'extrêmes  $a$  et  $b$ . On a vu dans la section précédente comment utiliser le changement de variable pour trouver une primitive de  $f$ . Il faut maintenant comprendre comment appliquer ce qu'on a vu dans le calcul d'une intégrale.

En effet nous avons le choix entre

- trouver une primitive de  $f$  comme dans la section précédente et appliquer ensuite la formule du corolaire 14.288 ;
- écrire une intégrale pour la nouvelle variable  $u = u(x)$  sur l'intervalle entre  $u(a)$  et  $u(b)$ .

Nous allons voir ces deux méthodes dans des exemples.

#### Exemple 20.141.

Soit à calculer

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx. \quad (20.520)$$

Les primitives  $\int x\sqrt{1-x^2}dx$  ont été trouvés dans l'exemple 20.140. Une primitive est

$$F(x) = \int x\sqrt{1-x^2}dx = -\frac{(1-x^2)^{3/2}}{3}. \quad (20.521)$$

Nous pouvons maintenant calculer l'intégrale de  $x\sqrt{1-x^2}$  sur l'intervalle  $[1/3, 1/2]$  par la définition

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx = F\left(\frac{1}{2}\right) - F\left(\frac{1}{3}\right) = -\frac{\sqrt{3}}{8} + \frac{16\sqrt{2}}{81}. \quad (20.522)$$

△

#### Remarque 20.142.

Pour que le calcul d'intégrale donne quelque chose de sensé il faut absolument que la primitive soit écrite en tant que fonction de  $x$  et non comme fonction de  $u$ . La méthode que nous allons voir dans l'exemple suivant réduit grandement la probabilité d'oublier ce détail, d'où le fait qu'elle soit de loin la plus utilisée.

#### Exemple 20.143.

Calculons à nouveau

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx. \quad (20.523)$$

Cette fois nous allons toucher à l'intervalle d'intégration en même temps que faire le changement de variables. Nous savons déjà les substitutions

$$\begin{cases} u = 1 - x^2 & (20.524a) \\ du = -2xdx & (20.524b) \\ dx = \frac{du}{-2x}. & (20.524c) \end{cases}$$



En ce qui concerne les extrêmes d'intégration, si  $x = 1/3$  alors  $u = 1 - \frac{1}{9} = \frac{8}{9}$  et si  $x = \frac{1}{2}$  alors  $u = \frac{3}{4}$ . Nous avons donc encore les substitutions suivantes :

$$\begin{cases} x = 1/3 \rightarrow u = 8/9 \\ x = 1/2 \rightarrow u = 3/4 \end{cases} \quad (20.525a)$$

$$(20.525b)$$

Le calcul est alors

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx = -\frac{1}{2} \int_{8/9}^{3/4} \sqrt{u}du = -\frac{1}{2} \left[ \frac{u^{3/2}}{3/2} \right]_{8/9}^{3/4} = -\frac{\sqrt{3}}{8} + \frac{16\sqrt{2}}{81}. \quad (20.526)$$

Attention : la dernière égalité n'est pas immédiate ; elle demande quelques calculs et une bonne utilisation des règles de puissances.  $\triangle$

La deuxième méthode est plus utilisée et, avec un peu d'exercice, plus rapide à mettre en place que la première.

Jusqu'à présent nous avons utilisé des changements de variables dans lesquels nous exprimions  $u$  en termes de  $x$ . Comme le montre l'exemple suivant, il est parfois fructueux d'utiliser le changement de variable dans le sens inverse : avec  $x$  exprimé en termes d'un paramètre.

exemplepassagepolaires

### Exemple 20.144.

À calculer :

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx. \quad (20.527)$$

Nous posons  $x = \sin(\theta)$  parce que nous savons que  $1 - \sin^2(\theta) = \cos^2(\theta)$  ; nous espérons que le changement de variables simplifie l'expression<sup>28</sup>. Les substitutions à faire dans l'intégrale sont :

$$\begin{cases} x = \sin(\theta) \\ dx = \cos(\theta)d\theta, \end{cases} \quad (20.528a)$$

$$(20.528b)$$

et en ce qui concerne les bornes, si  $x = 1/2$  alors  $\sin(\theta) = \frac{1}{2}$ , c'est-à-dire  $\theta = \frac{\pi}{6}$ . Si  $x = \sqrt{3}/2$  alors  $\theta = \frac{\pi}{3}$ . Donc

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx = \int_{\pi/6}^{\pi/3} \sqrt{1-\sin^2(\theta)} \cos(\theta)d\theta. \quad (20.529)$$

Nous avons  $1 - \sin^2(\theta) = \cos^2(\theta)$  et vu que  $\theta \in [\frac{\pi}{6}, \frac{\pi}{3}]$  nous avons toujours  $\cos(\theta) > 0$ , ce qui donne  $\sqrt{\cos^2(\theta)} = \cos(\theta)$ . Nous devons donc calculer

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta. \quad (20.530)$$

Pour celle-là, il faut utiliser une formule de trigonométrie<sup>29</sup> :

$$\cos^2(\theta) = \frac{1 + \cos(2\theta)}{2}. \quad (20.531)$$

Donc

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta = \int_{\pi/6}^{\pi/3} \frac{1 + \cos(2\theta)}{2}d\theta = \left[ \frac{\theta}{2} \right]_{\pi/6}^{\pi/3} + \int_{\pi/6}^{\pi/3} \frac{\cos(2\theta)}{2}d\theta, \quad (20.532)$$

Pour calculer proprement la dernière intégrale nous effectuons un autre changement de variable (facile) en posant  $t = 2\theta$ ,  $dt = 2d\theta$ ,  $t(\pi/6) = \pi/3$  et  $t(\pi/3) = 2\pi/3$ , nous avons alors

$$\int_{\pi/6}^{\pi/3} \cos^2(\theta)d\theta = \left[ \frac{\theta}{2} \right]_{\pi/6}^{\pi/3} + \int_{\pi/3}^{2\pi/3} \frac{\cos(t)}{4}dt = \left[ \frac{\theta}{2} \right]_{\pi/6}^{\pi/3} = \frac{\pi}{6} - \frac{\pi}{12} = \frac{\pi}{12}, \quad (20.533)$$

28. Lorsqu'on fait un changement de variables, il s'agit toujours d'espérer que l'expression se simplifie. Il n'y a pas moyen de savoir a priori si tel changement de variable va être utile. Il faut essayer.

29. En fait, il y a moyen de terminer le calcul en intégrant deux fois par parties, mais c'est plus compliqué.

parce que  $\sin\left(\frac{2\pi}{3}\right) = \sin\left(\frac{\pi}{3}\right)$ . Au final,

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2} dx = \frac{\pi}{12}. \quad (20.534)$$

△

### 20.25.4 Intégrations des fractions rationnelles réduites

#### Définition 20.145.

Une **fraction rationnelle** est un quotient de deux polynômes à coefficients réels ou complexes.

Par exemple

$$\frac{x^5 + 7x^4 - \frac{x^3}{2} + x}{x^2 - 1} \quad (20.535)$$

est une fraction rationnelle.

Il sera expliqué dans le cours d'algèbre que toute fraction rationnelle peut être écrite sous forme d'une somme d'éléments simples, c'est-à-dire de fractions rationnelles d'un des deux types suivants :

$$\frac{\alpha}{(x-a)^m}, \quad \alpha, a \in \mathbb{R}, m \in \mathbb{N} \quad (20.536a)$$

$$\frac{\alpha x + \beta}{(x^2 + ax + b)^m}; \quad \alpha, \beta, a, b \in \mathbb{R}, m \in \mathbb{N}, a^2 - 4b < 0. \quad (20.536b)$$

Nous allons nous contenter de donner un exemple de chaque type.

- (1) En ce qui concerne le cas (20.536a) avec  $m = 1$ , nous avons par exemple

$$\int \frac{1}{x-3} dx = \ln(|x-3|) + C. \quad (20.537)$$

Si vous voulez en être tout à fait sûr, effectuez d'abord le changement de variables  $u = x - 3$  qui donne  $dx = du$ .

- (2) En ce qui concerne le cas (20.536a) avec  $m \neq 1$ , nous avons par exemple

$$\int \frac{1}{(x-1)^4} dx = -\frac{1}{3(x-1)^3} + C. \quad (20.538)$$

Encore une fois, pour s'en convaincre, utiliser le changement de variables  $u = x - 1$ ,  $dx = du$  :

$$\int \frac{1}{(x-1)^4} dx = \int \frac{1}{u^4} du = \int u^{-4} du = -\frac{u^{-3}}{3} + C = -\frac{1}{3} \frac{1}{(x-1)^3} + C. \quad (20.539)$$

- (3) En ce qui concerne le cas (20.536b) avec  $\alpha \neq 0$ , nous avons par exemple

$$\int \frac{x}{x^2 + 4} dx = \frac{1}{2} \ln(x^2 + 4) + C. \quad (20.540)$$

Pour ce faire, il faut faire le changement de variables  $u = x^2 + 4$ ,  $du = 2x dx$ ,  $dx = \frac{du}{2x}$  qui donne

$$\int \frac{x}{x^2 + 4} dx = \frac{1}{2} \int \frac{du}{u} = \frac{1}{2} \ln(|u|) + C = \frac{1}{2} \ln(|x^2 + 4|) + C. \quad (20.541)$$

Dans ce cas nous pouvons oublier d'écrire la valeur absolue dans le logarithme parce que de toutes façons,  $x^2 + 4$  est toujours positif.

- (4) En ce qui concerne le cas (20.536b) avec  $\alpha = 0$ , nous avons par exemple

$$\int \frac{dx}{x^2 + 4} = \frac{1}{4} \int \frac{dx}{(\frac{x}{2})^2 + 1} = \frac{1}{2} \arctan\left(\frac{x}{2}\right) + C. \quad (20.542)$$

où nous avons utilisé la primitive  $\int \frac{dx}{x^2+1} = \arctan(x)$  du tableau de la page 1955. Pour vous en convaincre vous pouvez faire la dernière étape avec le changement de variables  $u = x/2$ ,  $dx = 2du$ .

## 20.25.5 Quelques formules à connaître

## À retenir 20.146

$$\int (\alpha f(x) + \beta g(x)) dx = \alpha \int f(x) dx + \beta \int g(x) dx. \quad (20.543a)$$

$$\int f(x)g'(x) dx = f(x)g(x) - \int f'(x)g(x) dx. \quad (20.543b)$$

$$\int f'(u(x))u'(x) dx = \int f(t) dt, \quad \text{avec } t = u(x). \quad (20.543c)$$

$$\int \frac{f'(x)}{f(x)} dx = \log |f(x)| + C, \quad \text{c'est un cas particulier de la formule précédente.} \quad (20.543d)$$

## 20.25.6 Taylor avec reste intégral

THOooDGCJooXKmFTT

**Théorème 20.147** (Taylor, reste intégral).Soit une fonction  $f: [a, b] \rightarrow \mathbb{R}$  de classe  $C^{n+1}$ . Alors pour tout  $N \leq n$  nous avons

$$f(b) = f(a) + \sum_{k=1}^N \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{1}{N!} \int_a^b (b-t)^N f^{(N+1)}(t) dt. \quad (20.544) \quad \text{EQooSCKCooXcKzCc}$$

*Démonstration.* Notons que dans l'énoncé,  $n$  est fixé ; nous faisons une récurrence sur  $N$ . Ça ne change pas grand chose, mais il faut être conscient de ce qui est exactement dans l'hypothèse du théorème et ce qui est dans l'hypothèse de récurrence.

Bref,  $n$  est fixé, la fonction  $f$  est de classe  $C^{n+1}$  et nous vérifions d'abord la formule avec  $N = 1$ . À droite dans (20.544) nous avons

$$f(a) + f'(a)(b-a) + \int_a^b (b-t)f''(t)dt. \quad (20.545) \quad \text{EQooETPRooJHcQXh}$$

Nous évaluons l'intégrale à part en faisant une intégration par parties<sup>30</sup>. Il s'agit de poser

$$u = b - t \quad (20.546a)$$

$$v' = f'', \quad (20.546b)$$

de déduire

$$u' = -1 \quad (20.547a)$$

$$v = f' \quad (20.547b)$$

et d'écrire

$$\int_a^b (b-t)f''(t)dt = [(b-t)f'(t)]_a^b - \int_a^b (-1)f'(t)dt \quad (20.548a)$$

$$= -(b-a)f'(a) + \int_a^b f'(t)dt \quad (20.548b)$$

$$= -(b-a)f'(a) + f(b) - f(a). \quad (20.548c)$$

Dans le calcul nous avons utilisé le théorème fondamental du calcul intégral 14.288. En remettant ça dans (20.545) nous trouvons  $f(b)$  comme il se doit.

30. Proposition 20.133.

En ce qui concerne la récurrence, nous devons calculer

$$f(a) + \sum_{k=1}^{N+1} \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{1}{(N+1)!} \int_a^b (b-t)^{N+1} f^{(N+2)}(t) dt. \quad (20.549)$$

Ici encore, il s'agit de faire une intégration par partie, et sortir de la somme le terme  $k = N + 1$ . L'intégration par partie donne

$$\int_a^b (b-t)^{N+1} f^{(N+2)}(t) dt = -(b-a)^{N+1} f^{(N+1)}(a) + (N+1) \int_a^b (b-t)^N f^{(N+1)}(t) dt. \quad (20.550)$$

En remettant tout ensemble, il y a encore deux termes qui se simplifient, et des termes qui se remettent pour former la formule de récurrence. Bref, on obtient que (20.549) se réduit bien à  $f(b)$ .  $\square$

Cette formule avec reste intégral sert par exemple à prouver un encadrement pour  $\ln(2)$ , voir la proposition 20.151.

**Proposition 20.148** (Formule de Taylor avec reste intégral[658]).

Soit une fonction  $f: \mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^{m+1}$  sur un ouvert  $U$  contenant  $a \in \mathbb{R}$ . Alors pour tout  $k \leq m$  et pour tout  $x \in U$  nous avons

$$f(x) = f(a) + \sum_{l=1}^k \frac{f^{(l)}(a)}{l!} (x-a)^l + \frac{1}{k!} \int_a^x f^{(k+1)}(t) (x-t)^k dt \quad (20.551a)$$

$$= \sum_{l=0}^k \frac{f^{(l)}(a)}{l!} (x-a)^l + \frac{1}{k!} \int_a^x f^{(k+1)}(t) (x-t)^k dt \quad (20.551b)$$

$$(20.551c)$$

*Démonstration.* Nous faisons par récurrence sur  $k$ . Nous commençons par voir la formule pour  $k = 0$ . La formule se réduit à

$$f(x) = f(a) + \int_a^x f'(t) dt, \quad (20.552)$$

qui est correcte par le théorème 14.288 sous la forme (14.911).

La vraie récurrence maintenant. Nous supposons donc avoir la formule

$$f(x) = f(a) + \sum_{l=1}^k \frac{f^{(l)}(a)}{l!} (x-a)^l + \frac{1}{k!} \int_a^x f^{(k+1)}(t) (x-t)^k dt \quad (20.553)$$

valide pour un certain  $k \leq m$ , et nous allons prouver la formule pour  $k + 1$ . L'idée est d'effectuer l'intégrale dans (20.553) par partie (proposition 20.133) en posant

$$\begin{aligned} u &= f^{(k+1)} & dv &= (x-t)^k \\ du &= f^{(k+2)} dt & v &= -\frac{(x-t)^{k+1}}{k+1}. \end{aligned} \quad (20.554)$$

Nous avons

$$\int_a^x f^{(k+1)}(t) (x-t)^k = \left[ -f^{(k+1)}(t) \frac{(x-t)^{k+1}}{k+1} \right]_a^x - \int_a^x -\frac{(x-t)^{k+1}}{k+1} f^{(k+2)}(t) dt \quad (20.555a)$$

$$= f^{(k+1)}(a) \frac{(x-a)^{k+1}}{k+1} + \int_a^x \frac{f^{(k+2)}(t)}{k+1} (x-t)^{k+1} dt. \quad (20.555b)$$

En remettant dans (20.553), nous trouvons la réponse.  $\square$

THOooFKZZooAgecfp

**Théorème 20.149** (Formule de Taylor, reste intégral[127]).

Soit une fonction  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^{m+1}$  sur un ouvert convexe  $U$ . Soient  $a \in U$  et  $h \in \mathbb{R}^n$  tel que  $a + h \in U$ . Alors

$$f(a + h) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq m}} \frac{h^\alpha}{\alpha!} (\partial^\alpha f)(a) + (m+1) \sum_{|\alpha|=m+1} \frac{h^\alpha}{\alpha!} \int_0^1 (1-t)^m (\partial^\alpha f)(a + th) dt. \quad (20.556)$$

*Démonstration.* Étant donné que  $U$  est connexe et que  $a$  et  $a + h$  sont dans  $U$ , nous pouvons poser

$$\begin{aligned} g: [0, 1] &\rightarrow \mathbb{R} \\ t &\mapsto f(a + th). \end{aligned} \quad (20.557)$$

Cela est une fonction de classe  $C^{m+1}$  parce que  $f$  l'est (théorème 10.304). Nous appliquons à la fonction  $g$  la formule du corolaire 12.328 :

$$g'(t) = h \cdot (\nabla f)(a + th). \quad (20.558) \quad \text{EQooQDLSooGhrePQ}$$

- (1) **Une récurrence pour  $g^{(k)}$**  Par  $h \cdot \nabla$  nous entendons l'opérateur  $h \cdot \nabla = h_1 \partial_1 + \dots + h_n \partial_n$ . Nous prouvons maintenant par récurrence que

$$g^{(k)}(t) = [(h \cdot \nabla)^k f](a + th). \quad (20.559)$$

En ce qui concerne le cas  $k = 1$ , c'est une réécriture de la formule (20.558) :

$$g'(t) = h \cdot (\nabla f)(a + th) = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \cdot \begin{pmatrix} (\partial_1 f)(a + th) \\ \vdots \\ (\partial_n f)(a + th) \end{pmatrix} \quad (20.560a)$$

$$= \sum_i h_i (\partial_i f)(a + th) \quad (20.560b)$$

$$= \sum_i (h_i \partial_i f)(a + th) \quad \text{cf. justif.} \quad (20.560c) \quad \text{SUBEQooTRIooPHkKyE}$$

$$= \left( \sum_i h_i \partial_i f \right)(a + th) \quad \text{cf. justif.} \quad (20.560d) \quad \text{SUBEQooHXSJooRZHTPT}$$

$$= ((h \cdot \nabla) f)(a + th). \quad (20.560e)$$

Justifications.

- Pour (20.560c). Dans un cas il s'agit de calculer  $(\partial_i f)(a + th)$  et de multiplier le résultat par  $h_i$ , et dans l'autre cas, de considérer la fonction  $(h_i \partial_i f)$  et de l'appliquer à  $a + th$ .
- Pour (20.560d). Avant, on calculait  $(h_i \partial_i f)(a + th)$  et on sommait sur  $i$ , et maintenant on applique la fonction  $\sum_i h_i \partial_i f$  à  $a + th$ .

Nous y allons maintenant pour la récurrence. Nous posons

$$s(t) = ((h \cdot \nabla)^k f)(a + th), \quad (20.561)$$

et nous essayons de calculer  $s'(t)$ . Nous avons

$$s'(t) = h \cdot \left( \nabla (h \cdot \nabla)^k f \right)(a + th) \quad (20.562a)$$

$$= \sum_i h_i \text{proj}_i \left[ \left( \nabla (h \cdot \nabla)^k f \right)(a + th) \right] \quad (20.562b)$$

$$= \sum_i h_i (\partial_i (h \cdot \nabla)^k f)(a + th) \quad (20.562c)$$

$$= \left( (h \cdot \nabla)^{k+1} f \right)(a + th). \quad (20.562d)$$

(2) Expression pour  $g^{(k)}$  En mettant les bouts ensemble, nous avons

$$g^{(k)}(t) = (h \cdot \nabla)^k f(a + th). \quad (20.563)$$

Nous prouvons à présent que

$$(h \cdot \nabla)^k = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = k}} \frac{k!}{\alpha!} h^\alpha \partial^\alpha. \quad (20.564)$$

Pour cela nous nous souvenons du théorème multinomial 3.37. Il s'applique parce que  $f$  est de classe  $C^{m+1}$ , ce qui permet de permuter les dérivées<sup>31</sup>. Les « variables »  $h_i \partial_i$  sont donc commutatives. D'abord nous avons

$$(h \cdot \nabla)^k = \left( \sum_i h_i \partial_i \right)^k = (h_1 \partial_1 + \dots + h_n \partial_n)^k, \quad (20.565)$$

et ensuite nous appliquons le théorème multinomial avec  $x_i = h_i \partial_i$ .

En ce qui concerne  $g$ , nous avons

$$g^{(k)}(t) = (h \cdot \nabla)^k f(a + th) = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = k}} \frac{k!}{\alpha!} (h^\alpha \partial^\alpha f)(a + th). \quad (20.566)$$

Nous savons que  $g(1) = f(a + h)$  et  $g(0) = f(a)$ . Nous écrivons donc un développement de Taylor pour  $g(1)$  autour de  $t = 0$ . Il s'agit d'utiliser la proposition 20.148.

$$g(1) = \sum_{l=0}^m \frac{g^{(l)}(0)}{l!} 1^l + \frac{1}{m!} \int_0^1 g^{(m+1)}(t) (1-t)^m dt \quad (20.567a)$$

$$= \sum_{l=0}^m \frac{1}{l!} \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = l}} \frac{l!}{\alpha!} (h^\alpha \partial^\alpha f)(a) + \frac{1}{m!} \int_0^1 \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = m+1}} \frac{(m+1)!}{\alpha!} (h^\alpha \partial^\alpha f)(a + th) (1-t)^{m+1} dt \quad (20.567b)$$

$$= \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq m}} \frac{1}{\alpha!} h^\alpha (\partial^\alpha f)(a) + (m+1) \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = m+1}} \int_0^1 \frac{h^\alpha}{\alpha!} (1-t)^m (\partial^\alpha f)(a + th). \quad (20.567c)$$

□

PROPOJFSRooGJcLyv

**Proposition 20.150** (Taylor à l'ordre  $m = 1[1]$ ).

Soit une fonction  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  de classe  $C^2$  sur un ouvert convexe  $U$ . Soient  $a \in U$  et  $h \in \mathbb{R}^n$  tel que  $a + h \in U$ . Alors

$$f(a + h) = h \cdot (\nabla f)(a) + h \cdot Q(a, h)h \quad (20.568)$$

où  $Q(a, h): \mathbb{R}^n \rightarrow \mathbb{R}^n$  est l'opérateur linéaire dont les éléments de matrice sont donnés par

$$Q(a, h)_{kl} = \int_0^1 (1-t) (\partial_{kl} f)(a + th) dt. \quad (20.569)$$

*Démonstration.* Nous commençons par écrire la formule Taylor avec reste intégral du théorème 20.149 avec  $m = 1$  :

$$f(a + h) = \sum_k h_k (\partial_k f)(a) + 2 \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| = 2}} \frac{h^\alpha}{\alpha!} \int_0^1 (1-t) (\partial^\alpha f)(a + th) dt \quad (20.570a)$$

$$= h \cdot (\nabla f)(a) + \frac{1}{2} \sum_{k,l} 2h_k h_l \int_0^1 (1-t) (\partial_{kl} f)(a + th) dt \quad \text{cf. justif} \quad (20.570b)$$

$$= h \cdot (\nabla f)(a) + \sum_l h_l Q(a, h)h \quad (20.570c)$$

$$= h \cdot (\nabla f)(a) + h \cdot Q(a, h)h \quad (20.570d)$$

$$(20.570e)$$

31. Proposition 12.379.

Justification pour (20.570b). Nous avons utilisé la formule du lemme 3.38 ainsi que le fait que  $h^{e_k+e_l} = h_k h_l$ .  $\square$

### 20.25.7 Approximation de $\ln(2)$

PROPooHOMYooFckCU

**Proposition 20.151** (Approximation de  $\ln(2)$ [659]).

Pour tout  $n$  nous avons

$$\left| \ln(2) - \sum_{k=1}^n \frac{(-1)^{k+1}}{k} \right| \leq \frac{1}{n+1}. \quad (20.571)$$

*Démonstration.* Nous écrivons la formule de Taylor avec reste intégral du théorème 20.147 pour la fonction  $f = \ln$  et pour  $a = 1$  et  $b = x$ . Cela donne :

$$\ln(x) = \ln(1) + \sum_{k=1}^N \frac{\ln^{(k)}(1)}{k!} (x-1)^k + \frac{1}{N!} \int_1^x (x-t)^N \ln^{(N+1)}(t) dt. \quad (20.572)$$

Sachant que la dérivée du logarithme<sup>32</sup> est  $1/x$  et faisant une petite récurrence, pour  $k \geq 1$  nous avons

$$\ln^{(k)}(x) = \frac{(-1)^{k+1} (k-1)!}{x^k}. \quad (20.573)$$

En remplaçant,

$$\ln(x) = \sum_{k=1}^N \frac{(-1)^{k+1}}{k} (x-1)^k + \int_1^x \frac{(-1)^N (x-t)^N}{t^{N+1}} dt. \quad (20.574)$$

C'est le moment de poser  $x = 2$  et de faire les simplifications qui s'imposent,

$$\ln(2) = \sum_{k=1}^N \frac{(-1)^{k+1}}{k} + \int_1^2 \frac{(-1)^N (2-t)^N}{t^{N+1}} dt. \quad (20.575)$$

Nous déplaçons la somme à gauche, et nous prenons la valeur absolue des deux côtés :

$$\left| \ln(2) - \sum_{k=1}^N \frac{(-1)^{k+1}}{k} \right| = \left| \int_1^2 \frac{(-1)^N (2-t)^N}{t^{N+1}} dt \right| \quad (20.576a)$$

$$\leq \int_1^2 \frac{(2-t)^N}{t^{N+1}} dt \quad (20.576b) \quad \text{SUBEQooZSXFooYchIpd}$$

$$\leq \int_1^2 (2-t)^N dt. \quad (20.576c) \quad \text{SUBEQooHGCZLooGIhoVt}$$

Justifications :

— Pour 20.576b. Majoration en rentrant la valeur absolue dans l'intégrale, suppression de  $(-1)^N$ , et le fait que pour  $t \in [1, 2]$ ,  $2-t \geq 0$ .

— Pour 20.576c. Majoration en supprimant purement et simplement le dénominateur  $t^{N+1} \geq 1$ .

Ais-je vraiment besoin de vous dire que la dernière intégrale se calcule en posant le changement de variables<sup>33</sup>  $u = 2-t$  ? Le résultat est que

$$\int_1^2 (2-t)^N dt = \frac{1}{N+1}. \quad (20.577)$$

$\square$

32. Voir la proposition 15.101.

33. Proposition 20.138.

**Exemple 20.152** ([1]).

La convergence de l'encadrement (20.151) n'est pas terrible. Pour avoir une erreur de  $\frac{1}{10}$ , il faut

$$\frac{1}{10} = \frac{1}{n+1}, \quad (20.578)$$

ce qui demande  $n = 9$ . Ça reste jouable, même pour les jeunes d'aujourd'hui. Écrivons 9 termes :

$$|\ln(2) - 1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \frac{1}{6} - \frac{1}{7} + \frac{1}{8} - \frac{1}{9}| \leq \frac{1}{10}. \quad (20.579)$$

En calculant<sup>34</sup>,

$$|\ln(2) - \frac{1879}{2520}| \leq \frac{1}{10}. \quad (20.580)$$

Voici donc un bel encadrement

$$\frac{1879}{2520} - \frac{1}{10} \leq \ln(2) \leq \frac{1879}{2520} + \frac{1}{10}. \quad (20.581)$$

Pour avoir quelque chose avec des virgules, d'abord un peu de calcul mental donne

$$\frac{1879}{2520} \simeq 0.745634920634921. \quad (20.582)$$

Donc en majorant et minorant, disons, la troisième décimale<sup>35</sup>, on n'est pas moins précis que le  $\frac{1}{10}$ . On a

$$0.744 - \frac{1}{10} \leq \ln(2) \leq 0.746 + \frac{1}{10}. \quad (20.583)$$

Bref, on retient l'approximation

$$0.644 \leq \ln(2) \leq 0.846. \quad (20.584)$$

Pour la quantité de travail, avouez que ce n'est pas terrible comme résultat. Eh oui ; le calcul numérique c'est tout un métier ; il existe des méthodes nettement plus efficaces que ce que nous venons de faire.  $\triangle$

## 20.26 Constructions plus naïves de l'intégrale dans le cas réel

Les sections 14.2 et 14.8 ont donné une construction très complète de la mesure de Lebesgue, et nous avons défini la théorie de l'intégration sur un espace mesuré quelconque dans la définition 14.167.

Dans cette section nous allons donner différentes choses plus rapides qui servent souvent de définition dans les cours moins avancés.

### 20.26.1 Mesure de Lebesgue, version rapide

Nous construisons à présent la mesure de Lebesgue sur  $\mathbb{R}^n$ . Un **pavé** dans  $\mathbb{R}^n$  est un ensemble de la forme

$$B = \prod_{i=1}^n [a_i, b_i]; \quad (20.585)$$

le volume d'un tel pavé est défini par  $\text{Vol}(B) = \prod_i (b_i - a_i)$ . Soit maintenant  $A \subset \mathbb{R}^n$ . La **mesure externe** de  $A$  est le nombre

$$m^*(A) = \inf \left\{ \sum_{B \in \mathcal{F}} \text{Vol}(B) \text{ où } \mathcal{F} \text{ est un ensemble dénombrable de pavés dont l'union recouvre } A. \right\} \quad (20.586)$$

34. Moi j'ai utilisé Sage, mais si tu es au tableau, débrouille-toi.

35. Notez ici que nous utilisons le fait que la division euclidienne, elle, donne un encadrement pour les fractions. Pensez-y.



DefKTzOlyH

**Définition 20.153.**

Nous disons que  $A$  est **mesurable** au sens de Lebesgue si pour tout ensemble  $S \subset \mathbb{R}^n$  nous avons l'égalité

$$m^*(S) = m^*(A \cap S) + m^*(S \setminus A). \quad (20.587)$$

Dans ce cas nous disons que la mesure de Lebesgue de  $A$  est  $m(A) = m^*(A)$ .

PropNCMToWI

**Proposition 20.154.**

Deux fonctions continues égales presque partout pour la mesure de Lebesgue<sup>36</sup> sont égales.

*Démonstration.* Soient  $f$  et  $g$  deux fonctions continues telles que  $f(x) = g(x)$  pour presque tout  $x \in D$ . La fonction  $h = f - g$  est alors presque partout nulle et nous devons prouver qu'elle est nulle sur tout  $D$ . La fonction  $h$  est continue; si  $h(a) \neq 0$  pour un certain  $a \in D$  alors  $h$  est non nulle sur un ouvert autour de  $a$  par continuité et donc est non nulle sur un ensemble de mesure non nulle.  $\square$

**20.26.2 Pavés et subdivisions****Définition 20.155.**

Nous appelons **pavé** de  $\mathbb{R}^p$  toute partie de  $\mathbb{R}^p$  obtenue comme produit de  $p$  intervalles de  $\mathbb{R}$ . Plus explicitement, une partie  $R$  est un pavé de  $\mathbb{R}^p$  si il s'écrit sous la forme

$$R = \{(x_1, \dots, x_p) \in \mathbb{R}^p \mid x_i \in \mathcal{I}_i, i = 1, \dots, p\},$$

où  $\mathcal{I}_i$  est un intervalle de  $\mathbb{R}$  pour tout  $i = 1, \dots, p$ .

On appelle pavé fermé de  $\mathbb{R}^p$  le produit de  $p$  intervalles fermés

$$R = \prod_{i=1}^p [a_i, b_i].$$

On définit de même le pavé ouvert

$$S = \prod_{i=1}^p ]a_i, b_i[.$$

Un pavé  $R = \prod_{i=1}^p \mathcal{I}_i$  est dit borné si tous les intervalles  $\mathcal{I}_i$  sont bornés dans  $\mathbb{R}$ . Les pavés non bornés sont des produits d'intervalles où un (ou plusieurs) des intervalles n'est pas borné. Par exemple,

$$N = ]-\infty, 5] \times [0, 13].$$

L'espace  $\mathbb{R}^p$ , lui-même, est un pavé de  $\mathbb{R}^p$ .

**Définition 20.156.**

Une partie  $A$  de  $\mathbb{R}^p$  est dite **pavable** si il existe une famille finie de pavés  $R_j$ ,  $j = 1, \dots, n$ , et deux à deux disjoints tels que

$$A = \bigcup_{j=1}^n R_j.$$

Un exemple d'ensemble pavable dans  $\mathbb{R}^2$  est donné à la figure 20.15. Il existe beaucoup d'ensembles dans  $\mathbb{R}^2$  qui ne sont pas pavables, par exemple les ellipses.

Le complémentaire d'un pavé est un ensemble pavable et, en particulier, tout complémentaire d'un pavé borné est une réunion de pavés non bornés. Toute union finie et toute intersection d'ensemble pavables est pavable.

---

36. Définition 20.153.

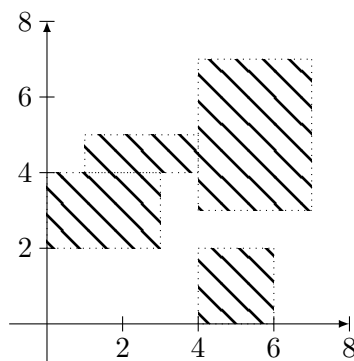


FIGURE 20.15: Un ensemble pavable.

LabelFigPolirettangolo

**Définition 20.157.**

Soit  $R$  un pavé borné de  $\mathbb{R}^p$ , pour fixer les idées on peut penser  $R = \prod_{i=1}^p [a_i, b_i]$ . On appelle **longueur** de l' $i$ -ème arête de  $R$  le nombre  $b_i - a_i$ . La **mesure  $p$ -dimensionnelle** de  $R$ ,  $m(R)$ , est le produit des longueurs

$$m(R) = \prod_{i=1}^p (b_i - a_i).$$

**Exemple 20.158.**

Dans  $\mathbb{R}^3$ , l'ensemble  $R = [-1, 1] \times [3, 4] \times [0, 2]$  est un pavé fermé de mesure

$$m(R) = (1 + 1) \cdot (4 - 3) \cdot (2 - 0) = 4.$$

Il s'agit du volume usuel du parallélépipède rectangle. △

**Exemple 20.159.**

L'ensemble  $R = ]-1, 1[ \times [3, 4] \times [0, 2]$  est un pavé de  $\mathbb{R}^3$ . Il n'est ni fermé ni ouvert, sa mesure est encore 4. △

Si  $R$  est un pavé non borné on peut encore définir sa mesure. La notion de mesure se généralise en deux étapes. D'abord on dit que la longueur d'une arête non bornée est  $\infty$ . Ensuite, on adopte la convention  $0 \cdot \infty = 0$ . Il faut remarquer que avec cette généralisation tout point et toute droite dans  $\mathbb{R}^2$  ont mesure nulle.

Afin de définir les intégrales, nous allons intensivement faire appel à la notion de subdivision d'intervalles, voir définition 21.2 et la discussion qui suit.

Lorsqu'on considère un pavé borné  $R = \prod_{i=1}^p \mathcal{I}_i$  de  $\mathbb{R}^p$ , on note  $\mathcal{S}_i$  l'ensemble des subdivisions de l'intervalle  $\mathcal{I}_i$ . La notion de subdivision se généralise au cas des pavés.

**Définition 20.160.**

Soit  $R$  un pavé fermé borné de  $\mathbb{R}^p$ , pour fixer les idées on peut penser à  $R = \prod_{i=1}^p [a_i, b_i]$ . On appelle **subdivision finie** de  $R$  les éléments de l'ensemble  $\mathcal{S} = \prod_{i=1}^p \mathcal{S}_i$ ,

$$\mathcal{S} = \left\{ (Y_1, \dots, Y_p) \mid Y_i = (y_{i,j})_{j=1}^{n_i} \in \mathcal{S}_i, i = 1, \dots, p \right\}.$$

On peut définir de même l'ensemble des subdivisions d'un pavé non borné.

Souvent, une subdivision d'un pavé  $R = \prod_{i=1}^p \mathcal{I}_i$  sera noté  $\sigma = (y_{i,j})_{j=1}^{n_i}$ . Dans cette notation, on sous-entend que pour chaque  $i$  fixé, les nombres  $y_{i,j}$  (il y en a  $n_i$ ) forment une subdivision de l'intervalle  $\mathcal{I}_i$ . Afin de vous familiariser avec ces notations, repérez bien tous les éléments de la figure 20.16.

**Définition 20.161.**

Si  $\sigma$  est une subdivision d'un pavé  $R$ , un **raffinement** de  $\sigma$  est une subdivision de  $R$  obtenue en fixant plus de points dans chaque intervalle.

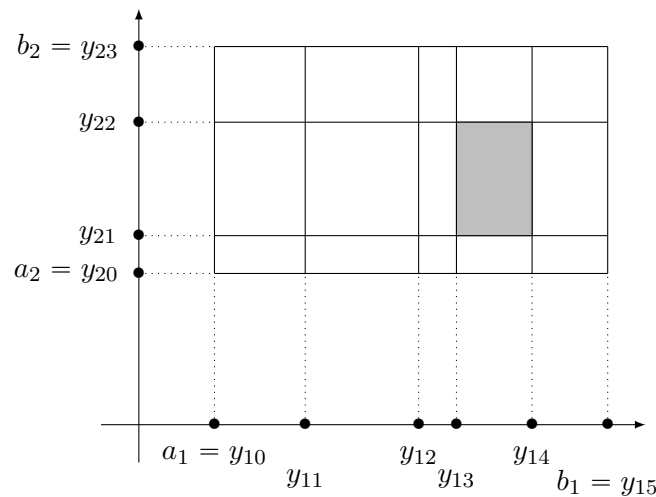


FIGURE 20.16: Une cellule d'une subdivision d'un pavé de  $\mathbb{R}^2$ . La cellule grisée est  $R_{(4,2)}$ .

La subdivision  $\sigma$  de  $R$  détermine  $n_1 n_2 \dots n_p$  pavés fermés de la forme

$$R_{(k_1, \dots, k_p)} = \{(x_1, \dots, x_p) \in \mathbb{R}^p \mid y_{i, k_{i-1}} \leq x_i \leq y_{i, k_i}\},$$

où  $k_i$  est dans  $\{1, \dots, n_i\}$  et  $i$  dans  $\{1, \dots, p\}$ . On les appelle **cellules** de  $\sigma$ . On remarque que les cellules de  $\sigma$  sont toujours deux à deux disjointes (sauf au plus sur leurs bords).

**Lemme 20.162.**

Soit  $R$  un pavé borné de  $\mathbb{R}^p$  et soit  $\sigma = (y_{i,j})_{j=1}^{n_i}$  une subdivision de  $R$ . On a

$$m(R) = \sum_{(k_1, \dots, k_p) \in K} m(R_{(k_1, \dots, k_p)}),$$

où  $K = \{1, \dots, n_1\} \times \{1, \dots, n_2\} \times \dots \times \{1, \dots, n_p\}$ .

Le lemme 20.162 suggère de définir la mesure d'un ensemble borné pavable  $P = \bigcup_{j=1}^n R_j$  comme la somme des mesures des pavés disjoints  $R_j$ ,  $j = 1, \dots, n$ .

**Définition 20.163.**

Une application  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  est dite **application en escalier** sur  $\mathbb{R}^p$  si

- $f$  est une application bornée,
- il existe une subdivision  $\sigma$  de  $\mathbb{R}^p$  telle que la restriction de  $f$  est une application constante sur toute cellule  $R_k$  de  $\sigma$

$$f|_{R_k} = C_k, \quad C_k \in \mathbb{R},$$

Une telle subdivision  $\sigma$  est dite **associée** à  $f$ .

**Exemple 20.164.**

La fonction  $f$  de  $\mathbb{R}^2$  dans  $\mathbb{R}$  définie par

$$f(x, y) = \begin{cases} 1 & \text{si } (x, y) \in [0, 3] \times [-1, 2], \\ 2 & \text{sinon.} \end{cases} \quad (20.588)$$

est une application en escalier. Exercice : donner une subdivision de  $\mathbb{R}^2$  associée à cette fonction.

△

**Exemple 20.165.**

La fonction  $f$  de  $\mathbb{R}^2$  dans  $\mathbb{R}$  définie par

$$f(x, y) = \begin{cases} \frac{1}{m^2 + n^2}, & \text{si } (x, y) \in [m, m+1] \times [n, n+1], \quad m, n \in \mathbb{N}_0, \\ 0, & \text{sinon} \end{cases} \quad (20.589)$$

est une application en escalier. Observez que, dans ce cas, il n'existe pas une subdivision finie de  $\mathbb{R}^2$  associée à  $f$ .  $\triangle$

**Remarque 20.166.**

Si la subdivision  $\sigma$  est associée à  $f$  alors tout raffinement de  $\sigma$  (c'est-à-dire, toute subdivision obtenue en fixant plus de points dans chaque intervalle) a la même propriété.

Si  $f$  et  $g$  sont deux applications en escalier sur  $R$  et  $\sigma_f$  et  $\sigma_g$  sont des subdivisions de  $R$  associées respectivement à  $f$  et  $g$ , alors on peut construire une troisième subdivision de  $R$  qui est associée à  $f$  et à  $g$  en même temps. Soient  $\sigma_f = (Y_1, \dots, Y_p)$  et  $\sigma_g = (Z_1, \dots, Z_p)$ , où  $Y_i = (y_{i,j})_{j=1}^{m_i}$  et  $Z_i = (z_{i,j})_{j=1}^{n_i}$  sont des subdivisions de l'intervalle  $[a_i, b_i]$ , pour  $i = 1, \dots, p$ . La subdivision de  $[a_i, b_i]$  obtenue par l'union de  $Y_i$  et  $Z_i$  est encore une subdivision finie, qu'on appellera  $\bar{Y}_i$ . La subdivision  $\bar{\sigma} = (\bar{Y}_1, \dots, \bar{Y}_p)$  de  $R$  est un raffinement de  $\sigma_f$  et de  $\sigma_g$ , donc elle est associée à la fois à  $f$  et à  $g$ .

Cela nous permet de prouver que si  $f$  et  $g$  sont des applications en escalier, alors  $f + g$ ,  $fg$ ,  $\min\{f, g\}$ ,  $\max\{f, g\}$  et  $|f|$  sont des applications en escalier.

### 20.26.3 Intégrale d'une fonction en escalier

**Définition 20.167.**

Soit  $f$  une fonction de  $\mathbb{R}^m$  dans  $\mathbb{R}^n$ . Le **support** de  $f$  est la fermeture de l'ensemble des points  $x$  tels que  $f(x) \neq 0$ .

**Définition 20.168.**

Une application en escalier  $f$  est dite **intégrable** si son support est compact.

Soit  $f$  une application en escalier sur  $\mathbb{R}^p$ . Soit  $\sigma$  une subdivision de  $\mathbb{R}^p$  associée à  $f$  et appelons  $R_k$  les cellules de  $\sigma$ , avec  $k = (k_1, \dots, k_p)$  dans  $K = \{1, \dots, n_1\} \times \{1, \dots, n_2\} \times \dots \times \{1, \dots, n_p\}$ . Alors

$$f|_{R_k} = C_k, \quad C_k \in \mathbb{R}.$$

**Définition 20.169.**

On définit l'**intégrale** de  $f$  sur  $\mathbb{R}^p$  par

$$\int_{\mathbb{R}^p} f dV = \sum_{k \in K} C_k m(R_k).$$

L'intégrale ainsi définie est un nombre réel. La proposition suivante nous dit que l'intégrale est « bien définie », au sens que sa valeur ne dépend pas de la subdivision associée à  $f$  qu'on utilise dans le calcul.

**Proposition 20.170.**

Soit  $f$  une application en escalier intégrable sur  $\mathbb{R}^p$ . Soient  $\sigma_1$  et  $\sigma_2$  deux subdivisions de  $\mathbb{R}^p$  associées à  $f$ . L'intégrale de  $f$  ne dépend pas de la subdivision choisie.

On ne donne pas une preuve complète de cette proposition. En fait elle est une conséquence de la formule de réduction introduite dans la suite de ce chapitre.

### 20.26.4 Intégrales partielles

Soit  $f$  de  $\mathbb{R}^p$  dans  $\mathbb{R}$  une fonction continue, nulle hors du pavé borné  $R$ . Posons  $R = \prod_{i=1}^p [a_i, b_i]$ , pour fixer les idées. Pour chaque  $i$  dans  $\{1, \dots, p\}$  fixé, on peut associer à  $f$  la fonction  $F_i$  de  $p-1$  variables définie par

$$F_i(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p) = \int_{a_i}^{b_i} f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_p) dy.$$

La fonction  $F_i$  est l'intégrale partielle de  $f$  par rapport à la  $i$ -ème variable. En particulier, si  $f(x_1, \dots, x_p) = g(x_i)h(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$  on obtient

$$F_i = \int_{a_i}^{b_i} g(y)h(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p) dy = h \cdot \int_{a_i}^{b_i} g dy.$$

La fonction d'une seule variable qu'on obtient à partir de  $f$  en fixant  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p$  et qui associe à  $x_i$  la valeur  $f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_p)$ , est appelée  $x_i$ -ème section de  $f$  en  $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p$ .

**Exemple 20.171.**

Soit  $f$  la fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$  définie par

$$f(x, y) = \begin{cases} x + 3y & \text{si } (x, y) \in [9, 10] \times ]\pi, 5] \\ 0 & \text{sinon.} \end{cases} \quad (20.590)$$

Les intégrales partielles de  $f$  sont

$$F_1(y) = \int_9^{10} x + 3y dx = \left[ \frac{x^2}{2} + 3xy \right]_{x=9}^{x=10} = \frac{19}{2} + 3y,$$

$$F_2(x) = \int_\pi^5 x + 3y dy = \left[ xy + \frac{3y^2}{2} \right]_{y=\pi}^{y=5} = x(5 - \pi) + \frac{3}{2}(25 - \pi^2).$$

△

### 20.26.5 Réduction d'une intégrale multiple

Soit  $R = [a, b] \times [c, d]$  un pavé fermé et borné de  $\mathbb{R}^2$  et soit  $f$  une application en escalier intégrable sur  $\mathbb{R}^2$  telle que le support de  $f$  soit contenu dans  $R$ . On considère la subdivision  $\sigma$  de  $R$  définie par les subdivisions

$$a = x_0 \leq x_1 \leq \dots \leq x_m = b,$$

$$c = y_0 \leq y_1 \leq \dots \leq y_n = d.$$

Les cellules de  $\sigma$  sont

$$R_{i,j} = [x_i, x_{i+1}] \times [y_j, y_{j+1}], \quad i = 0, \dots, m-1, \quad j = 0, \dots, n-1.$$

La mesure de  $R$  est la somme des mesures des  $R_{i,j}$

$$\begin{aligned} m(R) &= \sum_{(i,j) \in \{0, \dots, m-1\} \times \{0, \dots, n-1\}} m(R_{i,j}) = \\ &= \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} (x_{i+1} - x_i) \cdot (y_{j+1} - y_j) = \\ &= \sum_{i=0}^{m-1} (x_{i+1} - x_i) \cdot \sum_{j=0}^{n-1} (y_{j+1} - y_j) = \\ &= (b - a) \cdot (d - c). \end{aligned} \quad (20.591)$$

Si  $f$  est constante sur chaque cellule de  $\sigma$  on peut écrire  $f$  de la forme suivante

$$f(x, y) = \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} C_{i,j} \chi_{R_{i,j}}$$

où les  $C_{i,j}$  sont des constantes réelles et  $\chi_{R_{i,j}}$  est la **fonction caractéristique** de  $R_{i,j}$

$$\chi_{R_{i,j}}(x, y) = \begin{cases} 1, & \text{si } (x, y) \in R_{i,j}, \\ 0, & \text{sinon.} \end{cases} \quad (20.592)$$

Comme  $(x, y)$  est dans  $R_{i,j}$  si et seulement si  $x \in [x_i, x_{i+1}]$  et  $y \in [y_j, y_{j+1}]$ , on vérifie que la fonction  $\chi_{R_{i,j}}$  est égal au produit des fonctions caractéristiques des intervalles  $[x_i, x_{i+1}]$  et  $[y_j, y_{j+1}]$

$$\chi_{R_{i,j}}(x, y) = \chi_{[x_i, x_{i+1}]}(x) \cdot \chi_{[y_j, y_{j+1}]}(y).$$

On peut donc écrire la fonction  $f$  de la façon suivante

$$f(x, y) = \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} C_{i,j} \chi_{[x_i, x_{i+1}]}(x) \cdot \chi_{[y_j, y_{j+1}]}(y).$$

Comme on suppose que le support de  $f$  est une partie de  $R$ , l'intégrale de  $f$  sur  $\mathbb{R}^2$  est

$$\int_{\mathbb{R}^2} f dV = \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} C_{i,j} m(R_{i,j}) = \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} C_{i,j} (x_{i+1} - x_i) \cdot (y_{j+1} - y_j). \quad (20.593)$$

Cette intégrale peut être réduite à la composition de deux intégrales partielles. Il suffit de remarquer que la valeur de l'intégrale de la fonction caractéristique d'un intervalle est la longueur de l'intervalle,

$$\begin{aligned} C_{i,j}(x_{i+1} - x_i) \cdot (y_{j+1} - y_j) &= \\ &= C_{i,j} \left( \int_{x_i}^{x_{i+1}} \chi_{[x_i, x_{i+1}]}(x) dx \right) \cdot \left( \int_{y_j}^{y_{j+1}} \chi_{[y_j, y_{j+1}]}(y) dy \right) = \\ &= C_{i,j} \left( \int_a^b \chi_{[x_i, x_{i+1}]}(x) dx \right) \cdot \left( \int_c^d \chi_{[y_j, y_{j+1}]}(y) dy \right), \end{aligned} \quad (20.594)$$

et utiliser les propriétés de linéarité de l'intégrale

$$\begin{aligned} \int_{\mathbb{R}^2} f dV &= \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} C_{i,j} \left( \int_a^b \chi_{[x_i, x_{i+1}]}(x) dx \right) \cdot \left( \int_c^d \chi_{[y_j, y_{j+1}]}(y) dy \right) = \\ &= \int_c^d \int_a^b \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} C_{i,j} \chi_{[x_i, x_{i+1}]}(x) \cdot \chi_{[y_j, y_{j+1}]}(y) dx dy = \\ &= \int_c^d \int_a^b f dx dy. \end{aligned} \quad (20.595)$$

De même on obtient

$$\begin{aligned} \int_{\mathbb{R}^2} f dV &= \int_a^b \int_c^d \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} C_{i,j} \chi_{[x_i, x_{i+1}]}(x) \cdot \chi_{[y_j, y_{j+1}]}(y) dx dy = \\ &= \int_a^b \int_c^d f dx dy. \end{aligned} \quad (20.596)$$

En général, on prouve la proposition suivante

**Proposition 20.172.**

Soit  $f$  une application en escalier intégrable sur  $\mathbb{R}^p$  et soit  $R$  un pavé borné dans  $\mathbb{R}^p$  qui contient le support de  $f$ . Comme d'habitude, pour fixer les idées nous écrivons  $R = \prod_{i=1}^p [a_i, b_i]$ . Alors

$$\begin{aligned} \int_{\mathbb{R}^p} f(x_1, \dots, x_p) dV &= \int_{a_p}^{b_p} \int_{a_{p-1}}^{b_{p-1}} \cdots \int_{a_1}^{b_1} f(x_1, \dots, x_p) dx_1 \cdots dx_p = \\ &= \int_{a_{s_p}}^{b_{s_p}} \int_{a_{s_{p-1}}}^{b_{s_{p-1}}} \cdots \int_{a_{s_1}}^{b_{s_1}} f(x_1, \dots, x_p) dx_1 \cdots dx_p, \end{aligned} \quad (20.597)$$

pour toute permutation  $(s_1, \dots, s_p)$  de l'ensemble  $\{1, \dots, p\}$ .

### 20.26.6 Propriétés de l'intégrale

Soient  $f$  et  $g$  deux fonctions en escalier intégrables de  $\mathbb{R}^p$  dans  $\mathbb{R}$ , et soient  $a$  et  $b$  dans  $\mathbb{R}$ .

#### Linéarité de l'intégrale :

— Additivité :  $f + g$  est intégrable et

$$\int_{\mathbb{R}^p} (f + g) dV = \int_{\mathbb{R}^p} f dV + \int_{\mathbb{R}^p} g dV,$$

— Homogénéité :  $\lambda f$  est intégrable pour tout réel  $\lambda$

$$\int_{\mathbb{R}^p} \lambda f dV = \lambda \int_{\mathbb{R}^p} f dV,$$

**Monotonie** Si  $f \leq g$  alors

$$\int_{\mathbb{R}^p} f dV \leq \int_{\mathbb{R}^p} g dV,$$

#### Inégalité fondamentale

$$\left| \int_{\mathbb{R}^p} f dV \right| \leq \int_{\mathbb{R}^p} |f| dV.$$

Cette dernière inégalité s'obtient de la façon suivante :

$$\left| \int_{\mathbb{R}^p} f dV \right| = \left| \sum_{k \in K} C_k m(R_k) \right| \leq \sum_{k \in K} |C_k| m(R_k) = \int_{\mathbb{R}^p} |f| dV.$$

**Inégalité de Čebičeff** Si  $f$  est une application en escalier alors pour tout  $a > 0$  dans  $\mathbb{R}$  l'ensemble  $\{x \in \mathbb{R}^p : |f(x)| \geq a\}$  est pavable et borné, et l'inégalité suivante est satisfaite

$$m(\{x \in \mathbb{R}^p : |f(x)| \geq a\}) \leq \frac{1}{a} \int_{\mathbb{R}^p} |f| dV.$$

### 20.26.7 Intégrales multiples, cas général

Nous voulons généraliser la définition d'intégrale multiple au cas des domaines non pavables et de fonctions qui ne sont pas en escalier. Il y a plusieurs méthodes de le faire et ici on ne considère qu'une seule, introduite par Riemann.

#### Définition 20.173.

Soit  $f : \mathbb{R}^p \rightarrow \mathbb{R}$  une fonction.

- Pour toute application en escalier intégrable  $f_*$  telle que  $f_* \leq f$ , l'intégrale de  $f_*$  est dit une **somme inférieure** de  $f$ .
- Pour toute application en escalier intégrable  $f^*$  telle que  $f^* \geq f$ , l'intégrale de  $f^*$  est dit une **somme supérieure** de  $f$ .

Soient  $\Sigma_* f$  et  $\Sigma^* f$  les ensembles des sommes inférieures et supérieures de  $f$ . Grâce à la propriété de monotonie de l'intégrale on sait que si  $a$  est dans  $\Sigma_* f$  et  $b$  est dans  $\Sigma^* f$  alors  $a \leq b$ .

#### Définition 20.174.

La fonction  $f$  est intégrable (au sens de Riemann) si  $\Sigma_* f$  et  $\Sigma^* f$  ne sont pas vides et

$$\inf \Sigma^* f = I = \sup \Sigma_* f.$$

Dans ce cas, la valeur  $I$  est appelée intégrale de  $f$  sur  $\mathbb{R}^p$ .

#### Remarque 20.175.

Toute fonction intégrable est bornée et à support compact. En effet, si le support de la fonction n'est pas compact alors soit  $\Sigma_* f$  soit  $\Sigma^* f$  doit être vide !

L'intégrale qu'on vient de définir possède toutes les propriétés de l'intégrale pour les fonctions en escalier. Le produit de deux fonctions intégrables est intégrable.

Il y a des cas où l'intégrabilité d'une fonction n'est pas évidente. Cependant, dans la plupart des exercices et des exemples de ce cours, nous nous aidons avec le critère suivant

**Proposition 20.176.**

*Toute fonction continue à support compact est intégrable.*

Cette proposition n'est à priori pas étonnante, vu qu'une fonction continue sur un support compact est bornée (théorème de Weierstrass 7.196).

### 20.26.8 Réduction d'une intégrale multiple

On n'utilise jamais la définition pour calculer la valeur d'une intégrale multiple. La méthode plus efficace, en pratique, est de réduire l'intégrale à la composition de plusieurs intégrales d'une variable.

fub

**Théorème 20.177** (de Fubini).

*Soit  $f$  une fonction intégrable de  $\mathbb{R}^2$  dans  $\mathbb{R}$ . Si pour tout  $x$  dans  $\mathbb{R}$  la section  $f(x, \cdot)$  est intégrable par rapport à  $y$ , alors*

$$\int_{\mathbb{R}^2} f(x, y) dV = \int_{\mathbb{R}} \left( \int_{\mathbb{R}} f(x, y) dx \right) dy.$$

*De même, si pour tout  $y$  dans  $\mathbb{R}$  la section  $f(\cdot, y)$  est intégrable par rapport à  $x$ , alors*

$$\int_{\mathbb{R}^2} f(x, y) dV = \int_{\mathbb{R}} \left( \int_{\mathbb{R}} f(x, y) dy \right) dx.$$

En général, on ne peut pas dire que les sections d'une fonction intégrable sont intégrables, donc il faut vraiment se souvenir des hypothèses du théorème 20.177. En dimension plus haute, on a le même résultat

TheSectionINte

**Théorème 20.178.**

*Soit  $f$  une fonction intégrable de  $\mathbb{R}^p$  dans  $\mathbb{R}$ . Si pour tout  $(p-1)$ -uple  $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p)$  dans  $\mathbb{R}^{p-1}$  la section  $f(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_p)$  est intégrable par rapport à  $x_i$ , alors*

$$\int_{\mathbb{R}^p} f dV = \int_{\mathbb{R}} \left( \int_{\mathbb{R}^{p-1}} f dV \right) dx_i.$$

Si  $f$  est une fonction positive et intégrable de  $\mathbb{R}^2$  dans  $\mathbb{R}$  on peut interpréter l'intégrale de  $f$  comme le volume du solide au-dessous du graphe de  $f$ . Avec cette interprétation, l'intégrale partielle par rapport à  $x$  pour  $y = y_0$  fixé est l'aire de la tranche qu'on obtient en coupant le solide par le plan  $y = y_0$ .

**Exemple 20.179.**

Le premier exemple à faire est celui d'une fonction en escalier intégrable et positive. Soit  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  la fonction

$$f(x, y) = \begin{cases} 1 & \text{si } (x, y) \in R_1 = ]-1, 3] \times [4, 5] \\ 3 & \text{si } (x, y) \in R_2 = ]13, 15[ \times [0, 2[ \\ 0 & \text{dans les autres cas.} \end{cases} \quad (20.598)$$

L'intégrale de  $f$  sur  $\mathbb{R}^2$  est  $1 \cdot m(R_1) + 3 \cdot m(R_2) = 16$ . On voit tout de suite qu'il s'agit de la somme du volume des deux parallélépipèdes de hauteurs respectives 1 et 3 et bases  $R_1$  et  $R_2$ .  $\triangle$

**Exemple 20.180.**

On veut calculer le volume du solide  $S$ , borné par le paraboloïde elliptique  $x^2 + 2y^2 + z = 16$  et le plan  $x = 2, x = 0, y = 2, y = 0, z = 0$ . On observe que la portion de paraboloïde elliptique qui nous intéresse est le graphe de la fonction  $f(x, y) = 16 - x^2 - 2y^2$  pour  $(x, y)$  dans  $R = [0, 2] \times [0, 2]$ .



La fonction  $f$  est continue ainsi que ses sections, donc on peut appliquer le théorème 20.177 et décomposer l'intégrale double en deux intégrales simples :

$$\begin{aligned} \int_R 16 - x^2 - 2y^2 dV &= \int_0^2 \int_0^2 f(x, y) dx dy = \\ &= \int_0^2 \left[ (16 - 2y^2)x - \frac{x^3}{3} \right]_{x=0}^{x=2} dy = \\ &= \left[ \left( 32 - \frac{8}{3} \right) y - \frac{4y^3}{3} \right]_{y=0}^{y=2} = 64 - \frac{16 + 32}{3} = 48. \end{aligned} \quad (20.599)$$

Vérifiez, comme exercice, qu'on obtient le même résultat en intégrant d'abord par rapport à  $y$  et puis par rapport à  $x$ .  $\triangle$

### Exemple 20.181.

Dans les hypothèses du théorème 20.177 l'ordre des intégrations partielles ne change pas la valeur de l'intégrale. En fait, si les calculs sont faits par des êtres humains l'ordre d'intégration peut faire une certaine différence comme dans cet exemple. On veut évaluer la valeur de l'intégrale

$$\int_{\mathbb{R}^2} f(x, y) dV$$

où

$$f(x, y) = \begin{cases} y \sin(xy) & \text{si } (x, y) \in [1, 2] \times [0, \pi] \\ 0 & \text{sinon.} \end{cases} \quad (20.600)$$

Les deux sections de  $f(x, y) = y \sin(xy)$  sont continues. Si on intègre d'abord par rapport à  $y$  on obtient

$$- \int_1^2 \frac{\pi \cos(\pi x)}{x} dx + \int_1^2 \frac{\sin(\pi x)}{x^2} dx,$$

qui n'est pas du tout immédiat, alors que, si on intègre d'abord par rapport à  $x$  on obtient

$$\int_0^\pi \cos y - \cos(2y) dy.$$

$\triangle$

### 20.26.9 Intégrales sur des parties de $\mathbb{R}^2$

On veut évaluer l'intégrale de la fonction  $f(x, y) = \sqrt{1 - x^2}$  sur son domaine, la boule unité  $B((0, 0), 1)$ . La théorie introduite jusqu'ici n'est pas suffisante pour résoudre ce problème, parce que  $B((0, 0), 1)$  n'est pas pavable. Les parties bornées de  $\mathbb{R}^p$  sur lesquelles on peut intégrer des fonctions sont dites mesurables (au sens de Riemann) parce que, comme on verra dans la suite, la mesure d'une partie de  $\mathbb{R}^p$  est l'intégrale (si il existe) de sa fonction caractéristique.

On peut dire qu'une partie de  $\mathbb{R}^p$  est mesurable si son bord est «assez régulier». Dans  $\mathbb{R}^2$  il est suffisant que le bord de  $A$  soit une réunion finie de courbes paramétrées continues. En particulier, on est très souvent dans un des deux cas suivants

**Régions du premier type**  $A$  est borné et contenu entre les graphes de deux fonctions continues de  $x$

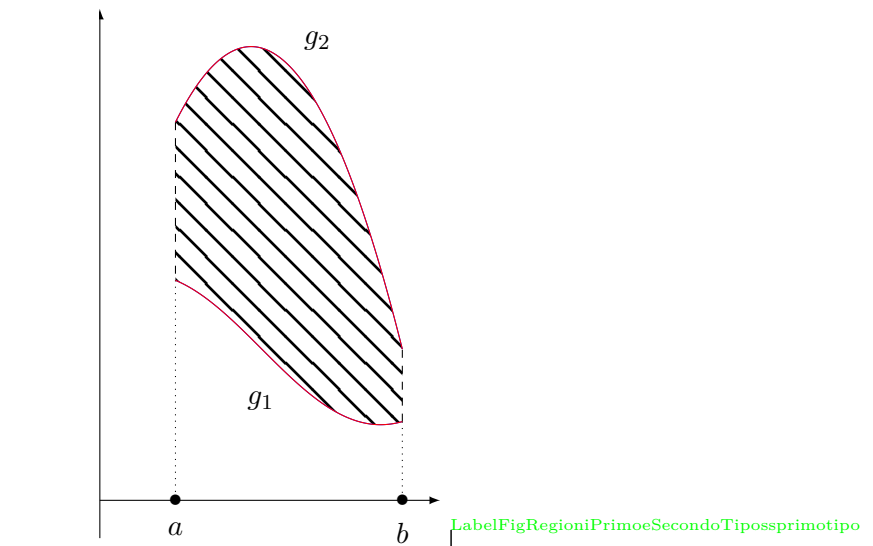
$$A = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, g_1(x) \leq y \leq g_2(x)\},$$

avec  $g_1$  et  $g_2$  continues.

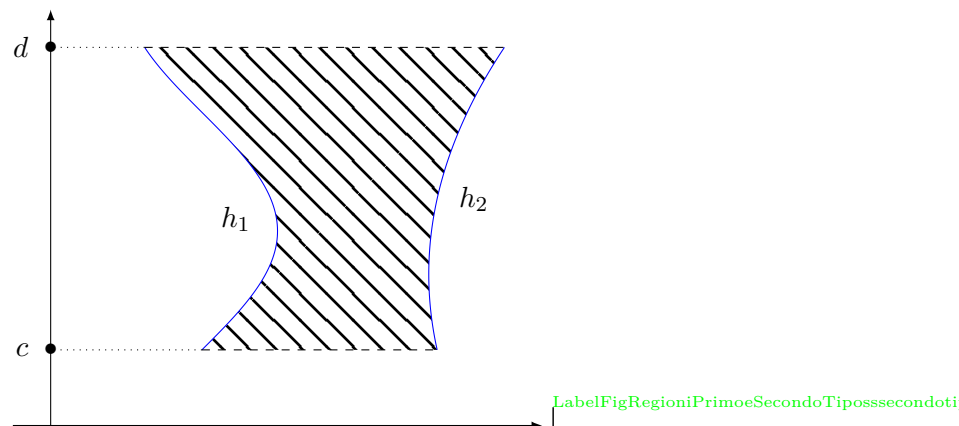
**Régions du deuxième type**  $A$  est borné et contenu entre les graphes de deux fonctions continues de  $y$

$$A = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d, h_1(y) \leq x \leq h_2(y)\},$$

avec  $h_1$  et  $h_2$  continues.



(a) Une region du premier type



(b) Une region du deuxième type

FIGURE 20.17: Régions du premier et du deuxième type

**Exemple 20.182.**

Il y a des régions qui sont des deux types au même temps, comme les boules centrées à l'origine, le triangle de sommets  $(0, 0)$ ,  $(0, a)$  et  $(b, 0)$ , ou la région  $C$  délimité par les courbes  $y = 2x$  et  $y = x^2$ . Cette dernière admet les représentations suivantes

$$C = \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1, x^2 \leq y \leq 2x\},$$

et

$$C = \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq 1, y/2 \leq x \leq \sqrt{y}\}.$$

△

**Définition 20.183.**

Soit  $f$  une fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$  dont le support  $A$  est une région du premier ou du deuxième type. On définit la fonction  $\bar{f}$  comme

$$\bar{f}(x, y) = \begin{cases} f(x, y), & \text{si } (x, y) \in A, \\ 0, & \text{sinon.} \end{cases} \quad (20.601)$$

La fonction  $f$  est dite **intégrable** si  $\bar{f}$  est intégrable, et la valeur de son intégrale est

$$\int_A f dV = \int_{\mathbb{R}^2} \bar{f} dV.$$

Une fonction continue définie sur une région du premier ou du deuxième type est toujours intégrable.

Pour fixer les idées on suppose ici que  $A$  est du premier type et contenue dans le pavé borné  $R = [a, b] \times [c, d]$ . En suivant la définition on obtient

$$\begin{aligned} \int_A f dV &= \int_{\mathbb{R}^2} \bar{f} dV \\ &= \int_a^b \int_c^d \bar{f} dy dx \\ &= \int_a^b \left( \int_c^{g_1(x)} \bar{f} dy + \int_{g_1(x)}^{g_2(x)} \bar{f} dy + \int_{g_2(x)}^d \bar{f} dy \right) dx \\ &= \int_a^b \int_{g_1(x)}^{g_2(x)} f dy dx. \end{aligned} \quad (20.602)$$

De même, si  $A$  est du deuxième type on obtient

$$\int_A f dV = \int_c^d \int_{h_1(y)}^{h_2(y)} f dx dy. \quad (20.603)$$

### Exemple 20.184.

On peut maintenant résoudre notre problème de départ, évaluer l'intégrale de la fonction  $f(x, y) = \sqrt{1 - x^2}$  sur  $B((0, 0), 1)$ . Nous choisissons de décrire la boule unité de  $\mathbb{R}^2$  comme une région du premier type :  $B((0, 0), 1) = \{(x, y) : x \in [-1, 1], -\sqrt{1 - x^2} \leq y \leq \sqrt{1 - x^2}\}$ .

$$I = \int_B \sqrt{1 - x^2} dV = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1 - x^2} dy dx \quad (20.604)$$

La première intégrale à effectuer, par rapport à  $y$ , est l'intégrale d'une fonction constante. Ne pas oublier que l'on intègre  $\sqrt{1 - x^2}$  par rapport à  $y$  ; c'est bien une constante et l'intégrale consiste seulement à multiplier par  $y$  :

$$I = \int_{-1}^1 \left[ y \sqrt{1 - x^2} \right]_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} dx = 2 \int_{-1}^1 (1 - x^2) dx. \quad (20.605)$$

Cela est à nouveau une intégrale simple à effectuer. Le résultat est

$$2 \int_{-1}^1 (1 - x^2) dx = 2 \left[ x - \frac{x^3}{3} \right]_{x=-1}^{x=1} = \frac{8}{3}. \quad (20.606)$$

△

### Remarque 20.185.

Toutes les techniques d'intégration à une variable restent valables. Par exemple, lorsqu'une des intégrales est l'intégrale d'une fonction impaire sur un intervalle symétrique par rapport à zéro, l'intégrale vaut zéro.

NORMooDSNXooPhyHkx

### 20.186.

Par le lemme 14.175 nous savons que la mesure d'une région bornée de  $\mathbb{R}^2$  est l'intégrale de sa fonction caractéristique, si elle existe.

La mesure d'une région bornée de  $\mathbb{R}^2$  est dite son **aire**, et celle d'une région bornée de  $\mathbb{R}^3$  est son **volume**. Voir aussi la remarque 21.10.

exint

### Exemple 20.187.

On veut calculer l'aire de la région de la figure 20.18 définie par

$$A = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, x^3 - 1 \leq y \leq x\}.$$

On considère l'intégrale

$$\int_{\mathbb{R}^2} \chi_A dV = \int_0^1 \int_{x^3+1}^x 1 dy dx = \int_0^1 -x^3 + x + 1 dx = -\frac{1}{4} + \frac{1}{2} + 1 = \frac{5}{4}.$$

△

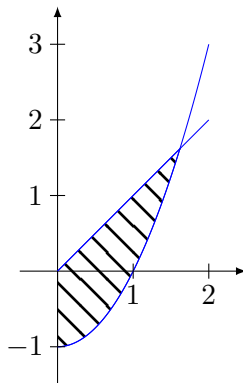


FIGURE 20.18: La région de l'exemple 20.187. Dans l'exemple, l'image est coupée en  $x = 1$ . Label: FigExampleIntegration

### Exemple 20.188.

Parfois la région sur laquelle on veut intégrer peut être décrite indifféremment de deux façons, mais la fonction à intégrer nous force à choisir un ordre particulier. Vérifiez que la fonction  $f(x, y) = \sin(y^2)$  sur la région triangulaire de sommets  $(0, 0)$ ,  $(0, 2)$ ,  $(2, 2)$  doit être intégrée d'abord par rapport à  $x$ . △

Si une région bornée n'est pas de premier ou de deuxième type on peut normalement la découper en morceaux plus faciles à décrire. On utilise alors la propriété suivante.

### Lemme 20.189.

Soit  $A$  un sous-ensemble borné de  $\mathbb{R}^2$  et soient  $B_1$  et  $B_2$  deux parties de  $A$  telles que  $B_1 \cap B_2 = \emptyset$  et  $B_1 \cup B_2 = A$ . Alors, pour toute fonction  $f$  intégrable sur  $A$  (et en particulier pour sa fonction caractéristique) on a

$$\int_A f dV = \int_{B_1} f dV + \int_{B_2} f dV.$$

exint2

### Exemple 20.190.

La région  $D$  que nous voyons sur la figure 20.19 est bornée par la parabole  $y^2 = 2x + 6$  et la droite  $y = x - 1$ . La région  $D$  est une région du deuxième type. Nous pouvons aussi la décrire comme l'union de deux régions du premier type  $D_1$  et  $D_2$ ,

$$D_1 = \{(x, y) : -3 \leq x \leq -1, -\sqrt{2x+6} \leq y \leq \sqrt{2x+6}\},$$

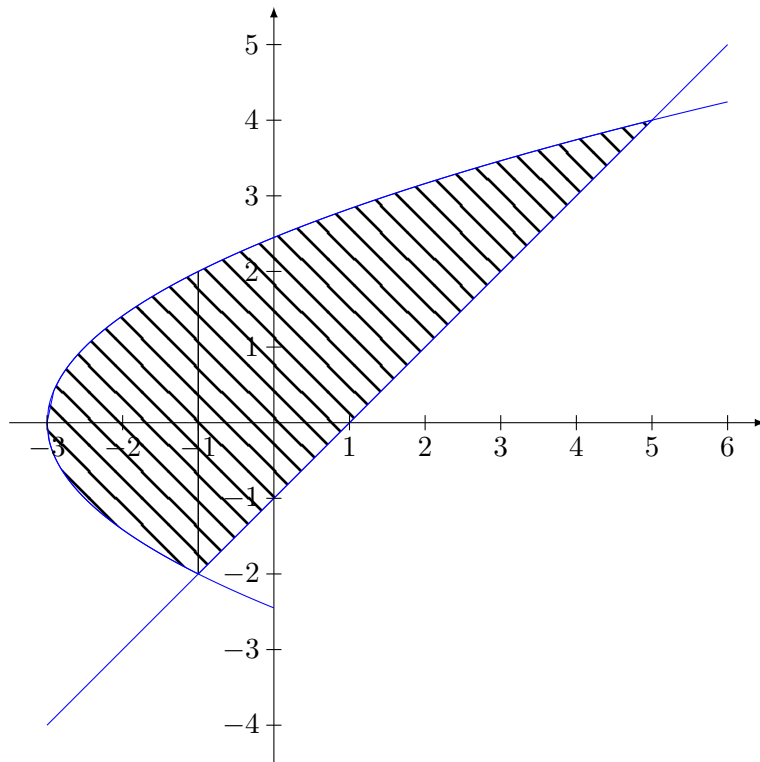
et

$$D_2 = \{(x, y) : -3 \leq x \leq -1, x - 1 \leq y \leq \sqrt{2x+6}\}.$$

△

## 20.26.10 Intégrales sur des parties de $\mathbb{R}^3$

Dans ces notes nous n'avons pas l'ambition de traiter d'une façon rigoureuse l'étude des ensembles mesurables de  $\mathbb{R}^3$ . Comme dans la section précédente on se limitera à considérer des cas particuliers.

FIGURE 20.19: La région  $D$  de l'exemple 20.190 LabelFigExampleIntegrationdeux**Définition 20.191.**

Soit  $E$  une région de  $\mathbb{R}^3$ . On dit que  $E$  est une **région solide de premier type** si  $E$  est contenue entre les graphes de deux fonctions continues de  $x$  et  $y$ .

$$E = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in A \subset \mathbb{R}^2, u_1(x, y) \leq z \leq u_2(x, y)\}.$$

Le sous-ensemble de  $A$  de  $\mathbb{R}^2$  qui apparaît dans la définition 20.191 est la projection (ou l'ombre) de  $E$  sur le plan  $x$ - $y$ .

**Exemple 20.192.**

La région  $E$  donnée par une portion de sphère collée à un cône est une région solide de premier type

$$E = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in \overline{B((0, 0), 1)}, \sqrt{x^2 + y^2} \leq z \leq \sqrt{1 - x^2 - y^2}\}. \quad (20.607)$$

L'ombre de  $E$  est la boule unité de  $\mathbb{R}^2$ . L'ensemble  $\sqrt{x^2 + y^2} \leq z$  est un cône posé sur sa pointe tandis que l'ensemble  $z \leq \sqrt{1 - x^2 - y^2}$  est la demi-sphère. L'ensemble  $E$  contient les points entre les deux, voir la figure 20.20.

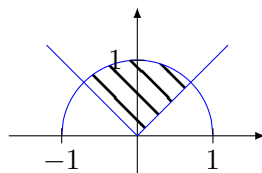


FIGURE 20.20: Il faut voir ça en trois dimensions.

LabelFigConeGlace

△

Si la fonction  $f$ , à intégrer sur  $E$ , et ses sections sont intégrables alors on peut réduire l'intégrale

$$\begin{aligned} \int_E f(x, y, z) dV &= \int_A \left( \int_{u_1(x, y)}^{u_2(x, y)} f(x, y, z) dz \right) dV = \\ &= \int_A (F(x, y, u_2(x, y)) - F(x, y, u_1(x, y))) dV, \end{aligned} \quad (20.608)$$

où  $F$  est une primitive de  $f$  par rapport à la variable  $z$ , c'est-à-dire en considérant  $x$  et  $y$  comme des constantes. Il faut ensuite évaluer la partie qui reste comme dans la section précédente. Comme le calcul des aires dans  $\mathbb{R}^2$ , le calcul des volumes dans  $\mathbb{R}^3$  est fait par des intégrales. En fait le **volume** d'une région solide dans  $\mathbb{R}^3$  est sa mesure.

**Définition 20.193.**

La mesure d'une région de  $\mathbb{R}^3$  est l'intégrale de sa fonction caractéristique.

Soit  $E$  une région solide du premier type, nous pouvons évaluer son volume par l'intégrale

$$\int_A (u_2(x, y) - u_1(x, y)) dV.$$

Parfois c'est plus intéressant de calculer le volume avec la formule de réduction contraire : l'intégrale double d'abord et puis l'intégrale simple par rapport à  $z$ . On parle alors de calcul de volume « par tranche ».

**Exemple 20.194.**

On veut calculer le volume de la boule de rayon  $a$ , centrée à l'origine  $B = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq a^2\}$ . On peut décrire  $B$  par

$$B = \left\{ (x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D_a, -\sqrt{a^2 - x^2 - y^2} \leq z \leq \sqrt{a^2 - x^2 - y^2} \right\},$$

où  $D_a$  est le disque de rayon  $a$  centré en  $(0, 0)$ , donc le volume  $B$  sera

$$2 \int_{D_a} \sqrt{a^2 - x^2 - y^2} dV.$$

Cette intégrale est un peu ennuyeuse à calculer. On peut simplifier le calcul en observant que pour  $\bar{z}$  fixé dans l'intervalle  $[-a, a]$  la section de la boule au niveau  $\bar{z}$  est un disque de rayon  $\sqrt{a^2 - \bar{z}^2}$ . L'aire d'un tel disque est  $\pi(a^2 - \bar{z}^2)$ . Si on réduit l'intégrale de volume de la façon

$$\int_B 1 dV = \int_{-a}^a \sqrt{a^2 - z^2} dz,$$

on obtient tout de suite la valeur cherchée : le volume de  $B$  est  $4/3\pi a^3$ . △

**Exemple 20.195.**

On calcule l'intégrale de  $f(x, y, z) = z$  sur la pyramide  $P$  bornée par les plans  $x = 0$ ,  $y = 0$ ,  $x + y + z = 1$ ,  $x + y + z/2 = 1$ . On remarque tout de suite que les plans  $x + y + z = 1$ ,  $x + y + z/2 = 1$  se coupent en la droite  $x + y = 1$ ,  $z = 0$  (on se souvient qu'une droite dans  $\mathbb{R}^3$ , c'est deux équations). Cela veut dire que la projection de  $P$  sur le plan  $x$ - $y$  est le triangle  $T$  borné par les droites  $x = z = 0$ ,  $y = z = 0$  et  $x + y = 1$ ,  $z = 0$ . On décrit donc  $P$  par

$$P = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in T, 1 - 2x - 2y \leq z \leq 1 - x - y\}$$

et  $T$  par

$$T = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, 0 \leq y \leq 1 - x\},$$

donc l'intégrale de  $f$  sur  $P$  est

$$\int_P f(x, y, z) dV = \int_0^1 \int_0^{1-x} \int_{1-2x-2y}^{1-x-y} z dz dy dx = -\frac{1}{24}.$$

Notez que lorsque  $x$  et  $y$  sont entre 0 et 1, nous avons bien  $1 - 2x - 2y < 1 - x - y$ , d'où le fait que nous mettons  $1 - 2x - 2y$  dans la borne inférieure de l'intégrale. △

De façon analogue on définit les régions solides du deuxième et du troisième type.

### 20.26.11 Intégrales de fonctions non bornées sur des ensembles non bornés

Soit  $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ , une fonction positive. Nous notons  $E_r = E \cap B(0, r)$ . On dit qu'elle est **intégrable** sur  $E \subset \mathbb{R}^n$  si

- (1) Pour chaque  $r > 0$ , la fonction  $f_r(x) = f(x)\mathbb{1}_{f < r}$  est intégrable sur  $E_r$ ;
- (2) la limite  $\lim_{r \rightarrow \infty} \int_{E_r} f_r$  est finie.

Dans ce cas, on pose

$$\int_E f = \lim_{r \rightarrow \infty} \int_{E_r} f_r. \quad (20.609)$$

#### Théorème 20.196.

TheFnTestIntnnBorn

Soit  $E$  mesurable dans  $\mathbb{R}^n$  et  $f: E \rightarrow \overline{\mathbb{R}}$ . Si  $f$  est mesurable et si il existe  $g: E \rightarrow \overline{\mathbb{R}}$  intégrable sur  $E$  telle que  $|f(x)| \leq g(x)$  pour tout  $x \in E$ , alors  $f$  est intégrable sur  $E$ .

Réciproquement, si  $f$  est intégrable sur  $E$ , alors  $f$  est mesurable.

LemTHBSEs

#### Lemme 20.197.

Si  $f$  est une fonction sur  $[a, \infty[$ , alors nous avons la formule

$$\lim_{b \rightarrow \infty} \int_a^b f(x) dx = \int_a^\infty f(x) dx \quad (20.610)$$

au sens où si un des deux membres existe, alors l'autre existe et est égal.

*Démonstration.* Supposons que le membre de gauche existe. Cela signifie que la fonction

$$\psi(x) = \int_a^x f \quad (20.611)$$

est bornée. Soit  $M$ , un majorant. Pour toute fonction simple  $\varphi$  dominée  $f$ , on a  $\int \varphi \leq M$ , donc l'ensemble sur lequel on prend le supremum pour calculer  $\int_a^\infty f$  est majoré par  $M$  et possède donc un supremum. Nous avons donc

$$\int_a^\infty f \leq \lim_{b \rightarrow \infty} \int_a^b f. \quad (20.612)$$

□

### 20.26.12 Lemme de Morse

LemNQAmCLO

#### Lemme 20.198 (Lemme de Morse).

Soit  $f \in C^3(\mathcal{U}, \mathbb{R})$  où  $\mathcal{U}$  est un ouvert de  $\mathbb{R}^n$  contenant 0. Nous supposons que

- (1)  $df_0 = 0$
- (2) La matrice  $H_{kl} = (\partial_{kl} f)(0)$  est non dégénérée et de signature  $(p, n - p)$ .

Alors il existe un  $C^1$ -difféomorphisme  $\varphi$  entre deux voisinages de 0 dans  $\mathbb{R}^n$  tel que

- (1)  $\varphi(0) = 0$ ,

- (2) si  $\varphi(x) = u$  alors

$$f(x) - f(0) = u_1^2 + \dots + u_p^2 - u_{p+1}^2 - \dots - u_n^2. \quad (20.613)$$

Une autre façon de dire est qu'il existe un  $C^1$ -difféomorphisme local  $\psi$  tel que

$$(f \circ \psi)(x) - f(0) = x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_n^2. \quad (20.614)$$

*Démonstration.* Nous utilisons le théorème de Taylor avec reste intégral sous la forme de la proposition 20.150 avec  $a = 0$  :

$$f(x) = x \cdot (\nabla f)(0) + x \cdot Q(x)x \quad (20.615)$$

avec

$$Q(x)_{kl} = \int_0^1 (1-t)(\partial_{kl}f)(tx)dt. \quad (20.616)$$

Par hypothèse sur  $df_0 = 0$ , le premier terme est nul. En ce qui concerne  $Q$ , nous avons

$$Q(0)_{kl} = \int_0^1 (1-t)(\partial_{kl}f)(0) = \frac{1}{2}H_{kl}. \quad (20.617)$$

Étant donné que  $f$  est de classe  $C^2$ , le théorème de Schwarz 12.370 dit que  $H$  est une matrice symétrique. Elle est également inversible par hypothèse.

À partir de là, le lemme 17.118 donne un voisinage  $V$  de  $Q(0)$  dans l'ensemble  $S_n$  des matrices symétriques et une application  $\phi$  de classe  $C^1$

$$\phi: V \rightarrow \text{GL}(n, \mathbb{R}) \quad (20.618)$$

telle que pour tout  $A \in V$ ,

$$\phi(A)^t Q(0) \phi(A) = A. \quad (20.619)$$

Si on pose  $M = \phi \circ Q$ , et si  $x$  est dans un voisinage de zéro,  $Q$  étant continue nous avons  $Q(x) \in V$  et donc

$$Q(x) = M(x)^t Q(0) M(x). \quad (20.620)$$

Notons que l'application  $M: \mathbb{R} \rightarrow \text{GL}(n, \mathbb{R})$  est de classe  $C^1$  parce que  $Q$  et  $\phi$  le sont.

Nous avons

$$f(x) - f(0) = x \cdot Q(x)x \quad (20.621a)$$

$$= x \cdot M(x)^t Q(0) M(x)x \quad (20.621b)$$

$$= M(x)x \cdot Q(0) M(x)x \quad \text{cf. justif} \quad (20.621c)$$

$$= y(x) \cdot Q(0)y(x) \quad \text{cf. justif} \quad (20.621d)$$

$$(20.621e)$$

Justifications.

— Pour 20.621c. Voir la proposition 8.186 et le blabla 8.187.

— Pour 20.621d. Nous notons  $y(x) = M(x)x = (\phi \circ Q)(x)x$ , qui est encore une fonction de classe  $C^1$  parce que la multiplication est une application  $C^\infty$ .

D'un autre côté le théorème de Sylvester 8.250 nous donne une matrice inversible  $P$  telle que

$$Q(0) = P^t \begin{pmatrix} \mathbb{1}_p & \\ & -\mathbb{1}_{n-p} \end{pmatrix} P. \quad (20.622)$$

Et nous posons enfin  $u = \varphi(x) = Py(x)$  qui est toujours de classe  $C^1$  et qui donne

$$f(x) - f(0) = y \cdot Q(0)y \quad (20.623a)$$

$$= y \cdot P^t \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} Py \quad (20.623b)$$

$$= u \cdot \begin{pmatrix} \mathbb{1} & \\ & -\mathbb{1} \end{pmatrix} u \quad (20.623c)$$

$$= u_1^2 + \dots + u_p^2 - u_{p+1}^2 - \dots - u_n^2. \quad (20.623d)$$

Nous devons maintenant montrer que, quitte à réduire son domaine à un ouvert plus petit,  $\varphi$  est un  $C^1$ -difféomorphisme. Dans la chaîne qui donne  $\varphi$ , seule l'application

$$\begin{aligned} g: U \subset \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\mapsto M(x)x \end{aligned} \quad (20.624)$$



est sujette à caution. Nous allons appliquer le théorème d'inversion locale. Nous savons que  $g$  est de classe  $C^1$  et donc différentiable ; calculons la différentielle en utilisant la formule (12.715) :

$$dg_0(x) = \frac{d}{dt} [g(tx)]_{t=0} = \frac{d}{dt} [tM(tx)x]_{t=0} = M(0)x. \quad (20.625)$$

Note que nous avons utilisé la règle de Leibniz pour la dérivée d'un produit, mais le second terme s'est annulé. Donc  $dg_0 = M(0) \in \text{GL}(n, \mathbb{R})$  et  $g$  est localement un  $C^1$ -difféomorphisme.

Il suffit de restreindre  $\varphi$  au domaine sur lequel  $g$  est un  $C^1$ -difféomorphisme pour que  $\varphi$  devienne lui-même un  $C^1$ -difféomorphisme. □

### Définition 20.199.

Un point  $a$  est un **point critique** de la fonction différentiable  $f$  si  $df_a = 0$ .

### Corolaire 20.200 ([660]).

Les points critiques non dégénérés d'une fonction  $C^3$  sont isolés.

*Démonstration.* Soit  $a$  un point critique non dégénéré. Par le lemme de Morse 20.198, il existe un  $C^1$ -difféomorphisme  $\psi$  et un entier  $p$  tel que

$$(f \circ \psi)(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_n^2 + f(a) \quad (20.626)$$

sur un voisinage  $\mathcal{U}$  de  $a$ . Vue la formule générale  $df_x(u) = \nabla f(x) \cdot u$ , si  $x$  est un point critique de  $f$ , alors  $\nabla f(x) = 0$ . Dans notre cas, les points critiques de  $f \circ \psi$  dans  $\mathcal{U}$  doivent vérifier  $x_i = 0$  pour tout  $i$ , et donc  $x = a$ .

Nous devons nous assurer que la fonction  $f$  elle-même n'a pas de points critiques dans  $\mathcal{U}$ . Pour cela nous utilisons la formule générale de dérivation de fonction composée :

$$\nabla(f \circ \psi)(x) = \sum_k \frac{\partial f}{\partial y_k}(g(x)) \nabla g_k(x). \quad (20.627)$$

Si  $\psi(x)$  est un point critique de  $f$ , alors le membre de droite est le vecteur nul parce que tous les  $\partial_k f(\psi(x))$  sont nuls. Par conséquent le membre de gauche est également nul, et  $x$  est un point critique de  $f \circ \psi$ . Or nous venons de voir que  $f \circ \psi$  n'a pas de points critiques dans  $\mathcal{U}$ .

Donc  $f$  n'a pas de points critiques dans un voisinage d'un point critique non dégénéré. □

## 20.27 Autres intégrales sympathiques

### 20.27.1 Intégrale de Wallis

#### Lemme 20.201 ([661, 1]).

Nous définissons les fonctions  $I_n$  par récurrence de la façon suivante :

$$\begin{cases} I_0(x) = x & (20.628a) \\ I_1(x) = \cos(x) & (20.628b) \end{cases}$$

$$I_n(x) = \frac{\cos(x) \sin^{n-1}(x)}{n} - \frac{n-1}{n} I_{n-2}(x). \quad (20.628c)$$

Pour chaque  $n$ , la fonction  $I_n$  est une primitive de  $\sin^n(x)$ .

De même en définissant  $I_n$  par

$$\begin{cases} I_0(x) = x & (20.629a) \\ I_1(x) = \sin(x) & (20.629b) \end{cases}$$

$$I_n(x) = \frac{\cos^{n-1}(x) \sin(x)}{n} + \frac{n-1}{n} I_{n-2}(x), \quad (20.629c)$$

nous avons des primitives de  $\cos^n(x)$ .

*Démonstration.* Nous faisons par récurrence. Les cas  $n = 0$  et  $n = 1$  sont immédiats. Pour  $I_n$  c'est pas trop compliqué non plus : il suffit de dériver en supposant que  $I'_{n-2}(x) = \sin^{n-2}(x)$ . Voici le calcul :

$$I'_n(x) = \frac{1}{n} \left( -\sin^n(x) + \cos(x)(n-1)\sin^{n-2}(x)\cos(x) \right) - \frac{n-1}{n} \sin^{n-2}(x) \quad (20.630a)$$

$$= \frac{1}{n} \sin^{n-2}(x) \left( -\sin^2(x) + (n-1)\cos^2(x) \right) - \frac{n-1}{n} \sin^{n-2}(x) \quad (20.630b)$$

$$= \frac{1}{n} \sin^{n-2}(x) \left( -\sin^2(x) + (n-1)\cos^2(x) - n + 1 \right) \quad (20.630c)$$

$$= \frac{1}{n} \sin^{n-2}(x) (n\cos^2(x) - n) \quad (20.630d)$$

$$= \frac{1}{n} \sin^{n-2}(x) n \sin^2(x) \quad (20.630e)$$

$$= \sin^n(x). \quad (20.630f)$$

Nous avons prouvé que la dérivée de  $I_n(x)$  est bien  $\sin^n(x)$ .

La formule (20.629) se démontre de la même façon. □

LEMooUOIBooLyMDft

**Lemme 20.202** ([662, 663]).

Nous posons

$$W_n = \int_0^{\pi/2} \cos^n(t) dt. \quad (20.631)$$

Alors :

(1) une formule de récurrence :

$$W_n = \frac{n-1}{n} W_{n-2}, \quad (20.632)$$

EQooILMZooBuzJnk

(2) et une formule un peu explicite :

$$W_{2n} = \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2}. \quad (20.633)$$

EQooUYIDooEpHCaP

*Démonstration.* Le nombre  $W_n$  est seulement la seconde intégrale du lemme 20.201, évaluée entre 0 et  $\pi/2$ . En partant donc de (20.629), nous avons

$$W_n = \int_0^{\pi/2} \cos^n(x) dx = \left[ \frac{\cos^{n-1}(x) \sin(x)}{n} \right]_0^{\pi/2} + \frac{n-1}{n} \int_0^{\pi/2} \cos^{n-2}(x) dx. \quad (20.634)$$

Le terme aux bords disparaît grâce aux valeurs trigonométriques remarquables<sup>37</sup>. Il reste immédiatement

$$W_n = \frac{n-1}{n} W_{n-2}. \quad (20.635)$$

À partir de là, nous démontrons (20.633) par récurrence. D'abord pour  $n = 0$ , c'est l'égalité  $W_0 = \pi/2$  qui est correcte parce que

$$\int_0^{\pi/2} \cos^0(x) dx = \int_0^{\pi/2} 1 dx. \quad (20.636)$$

Pour la récurrence elle-même,

$$W_{2(n+1)} = \frac{2n+1}{2(n+1)} W_{2n} \quad (20.637a)$$

$$= \frac{2n+1}{2n+2} \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2} \quad (20.637b)$$

$$\dots \text{pas mal de petits calculs} \dots \quad (20.637c)$$

$$= \frac{(2n+2)!}{((n+1)!2^{n+1})^2} \frac{\pi}{2}. \quad (20.637d)$$

Voilà. □

---

<sup>37</sup>. Par exemple la proposition 18.21(10).

Maintenant, la suite  $(W_n)$  se divise en ses termes pairs et ses termes impairs. Pour les pairs, nous avons une formule assez explicite donnée par le lemme 20.202. Pour les termes impairs, nous n'avons rien. Dans tous les cas, nous avons la formule de récurrence

$$W_n = \frac{n-1}{n} W_{n-2} \quad (20.638)$$

qui ne sert à rien pour déduire des choses sur les termes impairs à partir de ce que l'on sait des termes pairs.

Sommes-nous perdus ? Non. La situation se débloque grâce au lemme suivant.

LEM00ZFBVooQsOuOx

**Lemme 20.203.**

La suite donnée par

$$W_n = \int_0^{\pi/2} \cos^n(x) dx \quad (20.639)$$

est décroissante.

*Démonstration.* Vu que l'intégrale est sur  $[0, \pi/2]$ , le nombre  $\cos(x)$  prend ses valeurs dans  $[0, 1]$ . Nous avons donc

$$\cos^{n+1}(x) \leq \cos^n(x). \quad (20.640)$$

Les intégrales suivent les mêmes inégalités.  $\square$

Ce lemme permet de relancer le jeu parce que les termes impairs sont coincés entre les termes pairs, qui décroissent. Les termes impairs doivent donc décroître à la même vitesse. Le lemme suivant met cela en musique.

LEM00AXTEooLBXQuM

**Lemme 20.204.**

Nous posons

$$W_n = \int_0^{\pi/2} \cos^n(x) dx \quad (20.641)$$

La fonction  $\alpha: \mathbb{N} \rightarrow \mathbb{R}$  définie par  $W_n = \alpha(n)W_{n-1}$  vérifie  $\lim_{n \rightarrow \infty} \alpha(n) = 1$ .

*Démonstration.* Parce que nous en aurons besoin, nous triturons d'abord un peu la formule de récurrence (20.632). D'abord nous l'invertisons un peu pour avoir

$$W_n = \frac{n+2}{n} W_{n+2}, \quad (20.642)$$

et ensuite nous écrivons

$$W_{n-1} = \frac{n+1}{n} W_{n+1}. \quad (20.643)$$

EQooYINNooBUKUYB

Ne vous posez pas de questions, ça va être utile. Le fait que la suite soit décroissante (lemme 20.203) nous permet d'écrire

$$W_{n+1} \leq W_n \leq W_{n-1}. \quad (20.644)$$

En y remplaçant  $W_n$  par  $\alpha(n)W_{n+1}$  et  $W_{n-1}$  par (20.643),

$$W_{n+1} \leq \alpha(n)W_{n+1} \leq \frac{n+1}{n} W_{n+1}. \quad (20.645)$$

Nous simplifions par  $W_{n+1}$  et nous trouvons l'encadrement, valable pour tout  $n$  :

$$1 \leq \alpha(n) \leq \frac{n+1}{n}. \quad (20.646)$$

Nous en déduisons par la règle de l'étau que  $\alpha(n) \rightarrow 1$ .  $\square$

LEMooWQZAooOXAPQO

**Lemme 20.205.***Soit*

$$W_n = \int_0^{\pi/2} \cos^n(x) dx. \quad (20.647)$$

*Encore plusieurs choses à dire.*(1) *Pour tout  $n$  nous avons*

$$nW_nW_{n-1} = \frac{\pi}{2}. \quad (20.648)$$

EQooLOLFooMIwMXN

(2) *Nous avons l'équivalence de suites*<sup>38</sup>

$$W_n \sim \sqrt{\frac{\pi}{2n}}. \quad (20.649)$$

*Démonstration.* En deux parties(1) **La suite constante** Nous posons  $K_n = nW_nW_{n-1}$ . Grâce aux formules de récurrence (20.632) que nous écrivons sous la forme

$$W_n = \frac{n-1}{n} W_{n-2} \quad (20.650a)$$

$$W_{n+1} = \frac{n}{n+1} W_{n-1}, \quad (20.650b)$$

nous avons

$$K_{n+1} = (n+1)W_{n+1}W_n = (n+1)\frac{n}{n+1}\frac{n-1}{n}W_{n-1}W_{n-2} = (n-1)W_{n-1}W_{n-2} = K_{n-1}. \quad (20.651)$$

Nous avons montré que  $K_{n+1} = K_{n-1}$ .Il nous reste à prouver que  $K_1 = K_2 = \frac{\pi}{2}$ . Pour cela nous avons immédiatement  $W_0 = \frac{\pi}{2}$  ainsi que

$$W_1 = \int_0^{\pi/2} \cos(t) dt = 1. \quad (20.652)$$

Pour  $W_2$ , il ne faut pas calculer d'intégrales, mais seulement utiliser la formule (20.633). Nous trouvons vite  $W_2 = \frac{\pi}{4}$ . Donc

$$K_1 = W_1W_0 = \frac{\pi}{2} \quad (20.653)$$

et

$$K_2 = 2W_2W_1 = \frac{\pi}{2}. \quad (20.654)$$

(2) **L'équivalence de suites** Soit la fonction  $\alpha$  définie par  $W_n = \alpha(n)W_{n-1}$ . Le lemme 20.204 nous dit que  $\alpha(n) \rightarrow 1$ . Nous l'utilisons dans (20.648) :

$$\frac{\pi}{2} = nW_nW_{n-1} = n\alpha(n)W_{n-1}^2. \quad (20.655)$$

Donc

$$W_{n-1}^2 = \frac{\pi}{2n\alpha(n)}, \quad (20.656)$$

et

$$W_{n-1} \sqrt{\frac{1}{\alpha(n)}} \sqrt{\frac{\pi}{2n}}. \quad (20.657)$$

Vu que le coefficient  $\sqrt{1/\alpha(n)}$  tend vers 1 pour  $n \rightarrow \infty$ , nous avons l'équivalence demandée.

□

38. Définition 9.33.

### 20.27.2 Formule de Stirling

LEMooDXJOooOGFcIv

**Lemme 20.206** ([664]).

Si  $n \in \mathbb{N}$  nous avons la formule

$$\ln\left(1 + \frac{1}{n}\right) = 2 \sum_{k=0}^{\infty} \frac{1}{2k+1} \left(\frac{1}{2n+1}\right)^{2k+1} \quad (20.658)$$

*Démonstration.* Il s'agit d'utiliser astucieusement le développement de la proposition 15.110. Nous avons d'une part

$$\ln(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} \quad (20.659)$$

et d'autre part,

$$-\ln(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{k}. \quad (20.660)$$

Cela permet de calculer, en utilisant l'associativité de la série 39 EQooYVXHooNDetVx

$$\ln\left(\frac{1+t}{1-t}\right) = \ln(1+t) - \ln(1-t) \quad (20.661a)$$

$$= \sum_{k=1}^{\infty} \left( \frac{(-1)^{k+1} x^k}{k} + \frac{x^k}{k} \right) \quad (20.661b)$$

$$= 2 \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1} \quad (20.661c) \quad \text{SUBEQooIIOBooXYoHkK}$$

parce que tous les termes pairs s'annulent, tandis que les termes impairs sont doublés. Notez que la somme dans (20.661c) commence à  $k=0$ , contrairement aux autres qui commencent à 1.

Et là c'est l'astuce : on écrit l'égalité (20.661) avec le  $t$  qu'il faut pour que

$$\frac{1+t}{1-t} = 1 + \frac{1}{n}. \quad (20.662)$$

Nous posons donc  $t = \frac{1}{2n+1}$  nous avons

$$\ln\left(1 + \frac{1}{n}\right) = 2 \sum_{k=0}^{\infty} \frac{1}{2k+1} \left(\frac{1}{2n+1}\right)^{2k+1}, \quad (20.663)$$

ce qu'il fallait. □

LemCEoBqrP

**Lemme 20.207** (Formule de Stirling[665, 666, 662, 664, 667]).

Nous avons l'équivalence de suites 40

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}. \quad (20.664)$$

*Démonstration.* Nous posons

$$a_n = \frac{n!}{\sqrt{2n} \left(\frac{n}{e}\right)^n} \quad (20.665)$$

et  $b_n = \ln(a_n)$ .

---

39. Proposition 10.146.

40. Définition 9.33.

- (1) **Une formule pour  $b_n - b_{n+1}$**  Nous faisons un beau calcul qui utilise les formules de la proposition 15.94 ainsi que  $\ln(e) = 1$  : SUBEQSooZTABooBukFGM

$$b_n - b_{n+1} = \ln \left( \frac{a_n}{a_{n+1}} \right) \quad (20.666a)$$

$$= \ln \left( \frac{(n+1)^{n+1/2}}{n^{n+1/2}} \frac{1}{e} \right) \quad (20.666b)$$

$$= \ln \left( \left( \frac{n+1}{n} \right)^{n+1/2} \right) - 1 \quad (20.666c)$$

$$= \left( n + \frac{1}{2} \right) \ln \left( 1 + \frac{1}{n} \right) - 1. \quad (20.666d)$$

- (2) **La suite  $(b_n)$  est décroissante** Nous écrivons l'égalité (20.666) en utilisant le lemme 20.206 :

$$b_n - b_{n+1} = \left( n + \frac{1}{2} \right) \ln \left( 1 + \frac{1}{n} \right) - 1 \quad (20.667a)$$

$$= \frac{1}{2} (2n+1) 2 \sum_{k=0}^{\infty} \frac{1}{2k+1} \left( \frac{1}{2n+1} \right)^{2k+1} - 1 \quad (20.667b)$$

$$= (2n+1) \sum_{k=1}^{\infty} \frac{1}{2k+1} \left( \frac{1}{2n+1} \right)^{2k+1} \quad (20.667c)$$

$$= \sum_{k=1}^{\infty} \frac{1}{2k+1} \left( \frac{1}{2n+1} \right)^{2k}. \quad (20.667d)$$

Notez que le terme  $k=0$  s'est simplifié avec le  $-1$ . Vu que le tout est une somme de termes positifs, nous avons

$$b_n - b_{n+1} > 0 \quad (20.668)$$

et la suite est décroissante.

- (3) **Majoration pour  $b_n - b_{n+1}$**  Vu que  $\frac{1}{2k+1} < 1$ , nous pouvons majorer :

$$b_n - b_{n+1} \leq \sum_{k=1}^{\infty} \left( \frac{1}{(2n+1)^2} \right)^k. \quad (20.669)$$

Nous remarquons que cela est une série géométrique déjà traitée dans la proposition 10.159. Nous faisons un peu de calcul en partant de

$$\sum_{k=1}^{\infty} q^n = \frac{q}{1-q} \quad (20.670)$$

avec  $q = \frac{1}{(2n+1)^2}$ . Après quelques simplifications,

$$b_n - b_{n+1} \leq \frac{1}{4} \frac{1}{n(n+1)} \quad (20.671)$$

- (4) **Le coup de la somme télescopique** Nous avons

$$b_1 - b_n = (b_1 - b_2) + (b_2 - b_3) + \dots + (b_{n-1} - b_n). \quad (20.672)$$

Chacun de ces termes est majoré ; nous avons donc

$$b_1 - b_n < \frac{1}{4} \sum_{m=1}^{n-1} \frac{1}{m(m+1)} \leq \frac{1}{4} \sum_{m=1}^{\infty} \frac{1}{m(m+1)} = \frac{1}{4} \quad (20.673)$$

grâce au lemme 10.162 pour la dernière somme.

(5) La suite  $b_n$  est bornée vers le bas

Vu que  $b_1 = 1 - \frac{\ln(2)}{2}$ , nous avons

$$b_n > b_1 - \frac{1}{4} = \frac{3}{4} - \frac{\ln(2)}{2}. \quad (20.674)$$

En utilisant la majoration de l'exemple 20.152 nous trouvons

$$0.327 \leq b_n \leq 0.427. \quad (20.675)$$

Cet encadrement n'est pas très important. Le point est que la suite  $(b_n)$  soit bornée vers le bas ; savoir que la borne est strictement positive n'est pas indispensable.

(6) Une limite pour  $(a_n)$  La suite  $(b_n)$  est décroissante et bornée vers le bas, donc elle est convergente par le lemme 9.36. Vu que l'exponentielle est une fonction continue<sup>41</sup>, la suite  $a_n = e^{b_n}$  est également convergente.

Vu que  $\lim_{n \rightarrow \infty} b_n > 0$ , nous avons  $\lim_{n \rightarrow \infty} a_n = e^{\lim_{n \rightarrow \infty} b_n} > 1$ .

L'important est que nous sachions que  $(a_n)$  est une suite convergente. Nous notons  $L$  sa limite :

$$\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2n} \left(\frac{n}{e}\right)^n} = L. \quad (20.676)$$

Ou encore :

$$\lim_{n \rightarrow \infty} \frac{n!}{L\sqrt{2n} \left(\frac{n}{e}\right)^n} = 1. \quad (20.677)$$

Autrement dit, en posant

$$n! = \alpha(n)L\sqrt{2n} \left(\frac{n}{e}\right)^n, \quad (20.678)$$

nous avons  $\lim_{n \rightarrow \infty} \alpha(n) = 1$ .

(7) Introduction de Wallis Nous avons déjà parlé dans le lemme 20.202 du nombre

$$W_{2n} = \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2}. \quad (20.679)$$

Le lemme 20.205 implique que la suite  $I_n = W_{2n}$  est équivalente à  $\sqrt{\pi/4n}$ .

Cela étant dit, nous faisons un gros calcul en remplaçant les factorielles dans  $W_{2n}$  par la formule (20.678). Après pas mal de calculs<sup>42</sup>, nous trouvons

$$I_n = \frac{\alpha(2n)}{\alpha(n)^2} \frac{1}{\sqrt{n}} \frac{\pi}{2n}. \quad (20.680)$$

Nous avons donc

$$\sqrt{\frac{\pi}{4n}} \sim \frac{\alpha(2n)}{\alpha(n)^2} \frac{\pi}{\sqrt{n}} \frac{1}{2L}. \quad (20.681)$$

Il existe donc une fonction  $\beta(n)$  avec  $\beta(n) \rightarrow 1$  telle que

$$\frac{\sqrt{\pi}}{2} \frac{1}{\sqrt{n}} = \beta(n) \frac{\alpha(2n)}{\alpha(n)^2} \frac{\pi}{\sqrt{n}} \frac{1}{2L}. \quad (20.682)$$

En simplifiant par  $1/\sqrt{n}$ ,

$$\frac{\sqrt{\pi}}{2} = \beta(n) \frac{\alpha(2n)}{\alpha(n)^2} \frac{\pi}{2L}. \quad (20.683)$$

Nous prenons à présent la limite  $n \rightarrow \infty$  en nous rappelant que  $\alpha$  et  $\beta$  donnent 1. Après simplifications, nous trouvons

$$L = \sqrt{\pi}. \quad (20.684)$$

41. Parce qu'elle est dérivable, voir par exemple le théorème 15.84.

42. Si vous êtes en manque de papier de brouillon, c'est le moment de vous inquiéter.

(8) **La fin** Nous introduisons la valeur  $L = \sqrt{\pi}$  dans l'expression (20.678) de la factorielle :

$$n! = \alpha(n) \sqrt{2\pi n} \left(\frac{n}{e}\right)^n. \quad (20.685)$$

La dernière équation est exactement ce qui signifie l'équivalence de suite demandée.  $\square$

### 20.208.

J'ai déjà dit du mal de la formule d'Euler  $e^{i\pi} = -1$  dans (18.12). La formule de Stirling est plus intéressante parce qu'elle lie (réellement) les nombres  $e$  et  $\pi$ . Mais elle fait aussi intervenir le nombre  $\sqrt{2}$ , et la fonction factorielle.

La formule de Stirling est également nettement moins triviale à démontrer : la formule d'Euler ne fait intervenir de non trivial que les définition et propriétés des fonctions trigonométriques. Toutes ces propriétés sont également utilisées pour prouver la formule de Stirling.

### 20.27.3 La fonction sinus cardinal, intégrale de Dirichlet

#### Définition 20.209.

La fonction *sinus cardinal* est

$$f(t) = \begin{cases} 1 & \text{si } t = 0 \\ \frac{\sin(t)}{t} & \text{sinon.} \end{cases} \quad (20.686)$$

Elle sert à plein de choses. Entre autres, le lemme 20.212 montrera que la fonction  $x \mapsto |\sin(x)/x|$  a une intégrale sur  $\mathbb{R}$  qui vaut  $\infty$ . Cela nous permettra de donner un exemple d'une fonction dans  $L^1(\mathbb{R})$  dont la transformée de Fourier n'est pas dans  $L^1(\mathbb{R})$  (lemme 29.8).

### 20.210.

Le but que nous nous fixons maintenant est de prouver que

$$\int_0^\infty \frac{\sin(t)}{t} dt = \frac{\pi}{2}. \quad (20.687)$$

Un adage dit que si un théorème est trop long, c'est qu'il n'a pas assez de lemmes. Nous allons faire plein de lemmes.

LEMooMJFBooAjtNjV

#### Lemme 20.211.

La fonction *sinus cardinal* est continue.

*Démonstration.* Elle est continue en zéro parce que le lemme 18.246 nous donne

$$\lim_{t \rightarrow 0} \frac{\sin(t)}{t} = 1. \quad (20.688)$$

$\square$

Nous commençons par une mauvaise nouvelle.

LEMooEEWSooZwLSAP

#### Lemme 20.212 ([1]).

Nous avons

$$\int_{\mathbb{R}} \left| \frac{\sin(x)}{x} \right| dx = \infty. \quad (20.689)$$

*Démonstration.* Soit  $\delta > 0$  tel que sur l'intervalle  $[\frac{\pi}{2} - \delta, \frac{\pi}{2} + \delta]$ , nous ayons  $\sin(x) > 0.9$ <sup>43</sup>.

Les intervalles  $I_k = [\frac{\pi}{2} - \delta + 2k\pi, \frac{\pi}{2} + \delta + 2k\pi]$  sont disjoints et la fonction que nous intégrons est partout positive. Nous découpons

$$\mathbb{R} = C + \bigcup_{k=0}^{\infty} \left[ \frac{\pi}{2} - \delta + 2k\pi, \frac{\pi}{2} + \delta + 2k\pi \right] \quad (20.690)$$

43. Ça existe par une astucieuse combinaison du théorème 9.95 des valeurs intermédiaires, de la valeur remarquable  $\sin(\pi/2) = 1$  (de (18.42)) et du fait que  $\sin$  est continue (proposition 18.1).



où  $C$  est le complémentaire qu'il faut pour faire  $\mathbb{R}$ .

La  $\sigma$ -additivité de l'intégrale de Lebesgue (proposition 14.211) nous indique que

$$\int_{\mathbb{R}} \left| \frac{\sin(x)}{x} \right| dx = \int_C \left| \frac{\sin(x)}{x} \right| dx + \sum_{k=0}^{\infty} \int_{I_k} \left| \frac{\sin(x)}{x} \right| dx \quad (20.691)$$

Vu que tous les termes sont positifs, nous obtenons une majoration en en supprimant un. Allons-y :

$$\int_{\mathbb{R}} \left| \frac{\sin(x)}{x} \right| dx \geq \sum_{k=0}^{\infty} \int_{I_k} \left| \frac{\sin(x)}{x} \right| dx \quad (20.692a)$$

$$\geq \sum_{k=0}^{\infty} 0.9 \int_{I_k} \frac{1}{x} dx \quad (20.692b)$$

$$\geq 0.9 \sum_{k=0}^{\infty} 2\delta \frac{1}{\frac{\pi}{2} + \delta + 2k\pi} \quad (20.692c)$$

$$\geq 1.8\delta \sum_{k=0}^{\infty} \frac{1}{2\pi(k+1)} \quad (20.692d)$$

$$= \frac{1.8\delta}{2\pi} \sum_{k=0}^{\infty} \frac{1}{k+1}. \quad (20.692e)$$

Justifications :

- Pour (20.692c), nous avons majoré  $\frac{1}{x}$  par  $\frac{1}{\frac{\pi}{2} + \delta + 2k\pi}$  sur  $I_k$ .
- Pour (20.692d), nous avons dit que  $\frac{\pi}{2} + \delta < 2\pi$ .

La dernière somme dans (20.692) diverge.

Donc la fonction sinus cardinal n'est pas dans  $L^1(\mathbb{R})$ . □

La mauvaise nouvelle suivante en est un corolaire immédiat.

### Lemme 20.213.

La fonction  $t \mapsto \frac{\sin(t)}{t}$  n'est pas intégrable sur  $[0, \infty[$  au sens de Lebesgue.

*Démonstration.* Le lemme 20.212 nous dit que  $\int_0^\infty |f| = \infty$ . Dans ce cas,  $\int_0^\infty f$  n'existe pas par le lemme 14.188. □

Donc l'intégrale  $\int_0^\infty \frac{\sin(t)}{t} dt$  n'existe pas parce que la définition de l'intégrale de Lebesgue ne permet pas de profiter des compensations qui arrivent entre les valeurs positives et négatives.

Nous définissons donc

$$\int_0^\infty \frac{\sin(t)}{t} dt = \lim_{b \rightarrow \infty} \int_0^b \frac{\sin(t)}{t} dt. \quad (20.693)$$

Pour chaque  $b$ , l'intégrale existe sans problèmes (fonction continue sur le compact  $[0, b]$ ), et les compensations se font. Il n'est pas sans espoir que la limite (20.693) existe et vaille un nombre fini.

### Lemme 20.214 ([668]).

La limite

$$\lim_{b \rightarrow \infty} \int_0^b \frac{\sin(t)}{t} dt \quad (20.694)$$

existe dans  $\mathbb{R}$ .

*Démonstration.* En plusieurs parties.

(1) **Découpage**

Nous découpons l'intervalle  $[0, b]$  en morceaux du type  $[k\pi, (k+1)\pi]$  et un morceau restant lorsque  $b$  n'est pas un multiple de  $\pi$  :

$$[0, b] = \bigcup_{k=0}^{N(b)-1} [k\pi, (k+1)\pi] \cup [N(b)\pi, b] \quad (20.695)$$

où  $N(b)$  est un entier bien choisi<sup>44</sup>. En tout cas  $\lim_{b \rightarrow \infty} N(b) = \infty$ .

(2) **Majoration 1**

Pour chaque  $b \in \mathbb{R}^+$  nous avons

$$\int_0^b \frac{\sin(t)}{t} dt = \sum_{k=0}^{N(b)} \int_{[k\pi, (k+1)\pi]} \frac{\sin(t)}{t} dt + \int_{[(N(b)+1)\pi, b]} \frac{\sin(t)}{t} dt \quad (20.696)$$

Dans le dernier terme, nous majorons  $|\sin(t)| \leq 1$  et  $\frac{1}{t} \leq \frac{1}{(N(b)+1)\pi}$ . Cela donne

$$\left| \int_{[(N(b)+1)\pi, b]} \frac{\sin(t)}{t} dt \right| \leq \frac{b - (N(b)+1)\pi}{(N(b)+1)\pi} \leq \frac{1}{N(b)+1} \quad (20.697)$$

où nous avons encore majoré  $b \leq (N(b)+2)\pi$ .

(3) **Majoration 2**

En ce qui concerne les autres termes, sur l'intervalle  $[k\pi, (k+1)\pi]$ , nous avons  $\sin(t) = (-1)^k |\sin(t)|$ . Nous avons alors

$$\int_0^b \frac{\sin(t)}{t} dt = \sum_{k=0}^{N(b)} (-1)^k \int_{k\pi}^{(k+1)\pi} \frac{|\sin(t)|}{t} dt + \alpha(b) \quad (20.698)$$

où  $|\alpha(b)| \leq \frac{1}{N(b)+1}$  ; l'important est que  $\lim_{b \rightarrow \infty} \alpha(b) = 0$ .

(4) **Une suite alternée**

L'inégalité (20.698) nous incite à étudier la série  $\sum_{k=0}^{\infty} (-1)^k a_k$  en ayant posé

$$a_k = \int_{k\pi}^{(k+1)\pi} \frac{|\sin(t)|}{t} dt. \quad (20.699)$$

Nous montrons à présent que la suite  $(a_k)$  vérifie les conditions du critère des séries alternées 10.163.

D'abord,  $a_{k+1} \leq a_k$ . En effet en utilisant le changement de variables<sup>45</sup>  $u = t - \pi$ ,

$$a_{k+1} = \int_{(k+1)\pi}^{(k+2)\pi} \frac{|\sin(t)|}{t} dt = \int_{k\pi}^{(k+1)\pi} \frac{|\sin(u+\pi)|}{u+\pi} du = \int_{k\pi}^{(k+1)\pi} \frac{|\sin(u)|}{u+\pi} du < a_k. \quad (20.700)$$

Nous avons utilisé le fait que  $|\sin(u+\pi)| = |\sin(u)|$  pour tout  $u$ .

De plus, vu que  $|\sin(t)| \leq 1$  et que  $t \in [k\pi, (k+1)\pi]$ , nous avons

$$\int_{k\pi}^{(k+1)\pi} \frac{|\sin(t)|}{t} dt \leq \int_{k\pi}^{(k+1)\pi} \frac{1}{k\pi} dt = \frac{(k+1)\pi - k\pi}{k\pi} = \frac{1}{k}. \quad (20.701)$$

Donc  $a_k \leq \frac{1}{k} \rightarrow 0$ .

Le critère des séries alternées 10.163 nous dit que

$$\sum_{k=0}^{\infty} (-1)^k a_k < \infty. \quad (20.702)$$

44. Il me semble que le traitement de ce terme manque dans [668].

45. Le théorème 14.282 est toujours bon à citer.

(5) **Conclusion** Nous repartons de (20.698) :

$$\int_0^b \frac{\sin(t)}{t} dt = \sum_{k=0}^{N(b)} (-1)^k \int_{k\pi}^{(k+1)\pi} \frac{|\sin(t)|}{t} dt + \alpha(b). \quad (20.703)$$

Cette égalité est valable pour tout  $b \in \mathbb{R}^+$ . Le passage à la limite  $b \rightarrow 0$  à droite donne un nombre fini ; donc à gauche aussi, et nous avons prouvé que

$$\lim_{b \rightarrow \infty} \int_0^b \frac{\sin(t)}{t} dt < \infty. \quad (20.704)$$

□

Notre tâche n'est donc pas sans espoir. Au moins l'intégrale que nous cherchons à évaluer est finie.

LEMooARPIooDPSGwR

**Lemme 20.215.**

Soit  $x \in ]0, \infty[$ . L'intégrale

$$F(x) = \int_0^\infty e^{-tx} \frac{\sin(t)}{t} dt \quad (20.705)$$

existe au sens de Lebesgue usuel.

*Démonstration.* Vu qu'en  $t = 0$  nous avons  $\frac{\sin(t)}{t} = 1$ , il n'y a pas de problèmes de ce côté. Lorsque  $t > 1$  nous avons la majoration

$$|e^{-tx} \frac{\sin(t)}{t}| \leq |e^{-tx}|. \quad (20.706)$$

Lorsque  $t$  est assez grand, le lemme 15.115 nous donne aussi la majoration

$$|e^{-tx}| \leq \frac{1}{t^2}. \quad (20.707)$$

La proposition 14.301(2) implique que  $\int_1^\infty \frac{1}{t^2} < \infty$ . Et les majorations font que la proposition 14.176 nous donne le résultat. □

LEMooRDCSooBrWmep

**Lemme 20.216** ([668]).

Il existe une constante  $C \in \mathbb{R}$  telle que

$$I(x) = \int_0^\infty e^{-tx} \frac{\sin(t)}{t} dt = -\arctan(x) + C \quad (20.708)$$

pour tout  $x > 0$ .

*Démonstration.* En permutant dérivée et intégrale, nous allons prouver que  $I'(x) = -\frac{1}{1+x^2}$ .

(1) **Permuter** Nous posons

$$f(x, t) = e^{-tx} \frac{\sin(t)}{t}, \quad (20.709)$$

et nous vérifions les hypothèses du théorème 17.19.

(1a) Pour chaque  $x > 0$  fixé, la fonction  $t \mapsto f(x, t)$  est intégrable sur  $[0, \infty[$ , c'est le lemme 20.215.

(1b) Pour  $t > 0$  fixé, la fonction  $x \mapsto f(x, t)$  est dérivable.

(1c) Nous avons la dérivée partielle

$$\frac{\partial f}{\partial x}(x, t) = -e^{-tx} \sin(t) \quad (20.710)$$

qui vérifie

$$\left| \frac{\partial f}{\partial x}(x, t) \right| \leq e^{-tx}, \quad (20.711)$$

alors que la fonction  $t \mapsto e^{-tx}$  est intégrable sur  $[0, \infty[$ .

Nous pouvons donc dériver sous l'intégrale et obtenir

$$F'(x) = - \int_0^{\infty} e^{-xt} \sin(t) dt. \quad (20.712)$$

(2) Quelques intégrations par partie Nous posons

$$J(x) = \int_0^{\infty} e^{-xt} \sin(t) dt. \quad (20.713)$$

Cette intégrale est une intégrale de Lebesgue toute simple. En effet, posons provisoirement  $s(t) = e^{-xt} \sin(t)$ . La partie positive de  $s$  vérifie  $s_+(t) \leq e^{-xt}$  et la partie négative vérifie  $s_-(t) > -e^{-xt}$ . Les deux ont une intégrale qui converge. La définition 14.186 dit alors que  $J$  est une intégrale tout à fait propre. Elle serait d'ailleurs mieux écrite sous la forme

$$J(x) = \int_{[0, \infty]} e^{-xt} \sin(t) dt. \quad (20.714)$$

Bref. Tout ça pour dire que nous avons tout à fait le droit de la faire par parties<sup>46</sup>, en deux fois.

D'abord en posant  $u = e^{-xt}$  et  $v' = \sin(t)$  nous avons

$$J(x) = [-e^{-xt} \cos(t)]_{t=0}^{t=\infty} - \int_0^{\infty} (-x)e^{-xt} \cos(t) dt = 1 - x \int_0^{\infty} e^{-xt} \cos(t) dt. \quad (20.715)$$

Nous faisons l'intégrale encore par parties en posant  $u = e^{-xt}$  et  $v' = \cos(t)$  :

$$\int_0^{\infty} e^{-xt} \cos(t) dt = [e^{-xt} \sin(t)]_0^{\infty} - \int_0^{\infty} (-x)e^{-xt} \sin(t) dt = xJ(x). \quad (20.716)$$

Donc

$$J(x) = 1 - x \int_0^{\infty} e^{-xt} \cos(t) dt \quad (20.717a)$$

$$= 1 - x \int_0^{\infty} e^{-xt} \cos(t) dt \quad (20.717b)$$

$$= 1 - x \left( x \underbrace{\int_0^{\infty} e^{-xt} \sin(t) dt}_{J(x)} \right) \quad (20.717c)$$

$$= 1 - x^2 J(x). \quad (20.717d)$$

Voilà qui prouve que  $J(x) = \frac{1}{1+x^2}$ , et donc que

$$F'(x) = -J(x) = -\frac{1}{1+x^2}. \quad (20.718)$$

(3) Et enfin

Le théorème 18.39(2) nous dit que la dérivée de la fonction arctan est précisément  $1/(1+x)$ . Donc  $F$  et arctan ont la même dérivée (au signe près). Donc il existe  $C \in \mathbb{R}$  tel que

$$F(x) = -\arctan(x) + C \quad (20.719)$$

pour tout  $x > 0$ .

□

---

46. Proposition 20.133.

LEMooEOYHooVIMCCa

**Lemme 20.217.***Nous avons*

$$F(x) = \int_0^{\infty} e^{-xt} \frac{\sin(t)}{t} dt = -\arctan(x) + \frac{\pi}{2} \quad (20.720)$$

*pour tout  $x > 0$ .*

*Démonstration.* Le but de ce lemme est de fixer la constante laissée arbitraire dans le lemme 20.216. Nous savons qu'il existe  $C \in \mathbb{R}$  tel que

$$F(x) = -\arctan(x) + C. \quad (20.721)$$

EQooTZGXooUxfAjt

Le but est de prendre la limite  $x \rightarrow \infty$  des deux côtés.

Par le lemme 18.247, nous avons  $|\frac{\sin(t)}{t}| \leq 1$  sur  $[0, \infty[$ . Donc

$$F(x) \leq \int_0^{\infty} e^{-xt} dt = \left[ -\frac{1}{x} e^{-xt} \right]_{t=0}^{t=\infty} = \frac{1}{x}. \quad (20.722)$$

Vu que  $F(x) \leq \frac{1}{x}$  pour tout  $x$ , nous avons certainement  $\lim_{x \rightarrow \infty} F(x) = 0$ .

D'autre part,

$$\lim_{x \rightarrow \infty} \arctan(x) = \frac{\pi}{2}. \quad (20.723)$$

En passant à la limite dans (20.721), nous avons

$$0 = -\frac{\pi}{2} + C, \quad (20.724)$$

et donc  $C = \pi/2$ . □

Nous avons maintenant la formule

$$F(x) = \int_0^{\infty} e^{-tx} \frac{\sin(t)}{t} dt = -\arctan(x) + \frac{\pi}{2} \quad (20.725)$$

qui est valable pour tout  $x > 0$ .

Notre but sera de prendre la limite  $x \rightarrow 0$  des deux côtés. Vu que  $\arctan$  est continue, le membre de droite ne pose pas de problèmes et donne  $\pi/2$ . Pour le membre de gauche, il faut encore permuter une limite et une intégrale.

Pour la suite, nous allons étudier[668]

$$\int_0^{\infty} (1 - e^{-xt}) \frac{\sin(t)}{t} dt. \quad (20.726)$$

Cette intégrale n'existe pas au sens de Lebesgue et est définie par

$$L(x) = \lim_{b \rightarrow \infty} \int_0^b (1 - e^{-xt}) \frac{\sin(t)}{t} dt. \quad (20.727)$$

Rien n'indique cependant pour l'instant que cette limite existe.

LEMooZGODooLaBuHo

**Lemme 20.218 ([668]).**

*Soit  $x > 0$ . Nous posons*

$$L_k(x) = \int_{k\pi}^{(k+1)\pi} (1 - e^{-tx}) \frac{|\sin(t)|}{t} dt. \quad (20.728)$$

EQooJXWMooRbbCtt

*La suite  $(L_k(x))_{k \in \mathbb{N}}$  satisfait le critère des séries alternées<sup>47</sup>, c'est-à-dire que cette suite est positive, décroissante à limite nulle.*

---

47. Théorème 10.163.

*Démonstration.* Notons que chacune des intégrales  $L_k(x)$  est sans problèmes : fonction continue sur un compact. Trois éléments à prouver.

- (1) **Positive** Vu que dans toute notre histoire,  $x, t > 0$ , nous avons  $1 - e^{-tx} > 0$  et donc toute la fonction intégrée est positive.
- (2) **Tend vers zéro** Vu que  $1 - e^{-tx} < 1$ , nous avons

$$L_k(x) \leq \int_{k\pi}^{(k+1)\pi} \frac{1}{t} dt \leq \pi \frac{1}{k\pi} = \frac{1}{k}. \quad (20.729)$$

Donc  $\lim_{k \rightarrow \infty} L_k(x) = 0$ .

- (3) **Décroissante** Nous devons à présent prouver que  $L_k(x)$  est décroissante en  $k$  lorsque  $x$  est fixé.

Nous avons  $L_{k+1}(0) = L_k(0)$  pour tout  $k$ . Nous allons montrer que  $L'_{k+1}(x) < L'_k(x)$  pour tout  $x > 0$ . De cette façon nous aurons bien  $L_{k+1}(x) < L_k(x)$  pour tout  $k$  et  $x$ .

En permutant (encore) intégrale et dérivée,

$$L'_{k+1}(x) = \int_{(k+1)\pi}^{(k+2)\pi} e^{-tx} |\sin(t)| dt \quad (20.730a)$$

$$= \int_{k\pi}^{(k+1)\pi} e^{-(u+\pi)x} |\sin(u + \pi)| \quad (20.730b)$$

$$= e^{-\pi x} \int_{k\pi}^{(k+1)\pi} e^{-ux} |\sin(u)| du \quad (20.730c)$$

$$= e^{-\pi x} L'_k(x). \quad (20.730d)$$

Justifications :

- Pour (20.730a), permuter dérivée et intégrale ; je ne donne pas tout le détail. Ça a déjà été fait.
- Pour (20.730b), nous avons fait le changement de variables  $u = t - \pi$ .
- Pour (20.730c), nous avons utilisé le fait que  $|\sin(u + \pi)| = |\sin(u)|$  ainsi que  $e^{-(u+\pi)x} = e^{-ux} e^{-\pi x}$  par (12.1164).

Nous avons donc prouvé que

$$L'_{k+1}(x) = e^{-\pi x} L'_k(x) < L'_k(x). \quad (20.731)$$

□

LEM00SWFD00GLfwoD

### Lemme 20.219.

Pour chaque  $x > 0$ , nous avons la limite

$$\lim_{b \rightarrow \infty} \int_0^b (1 - e^{-tx}) \frac{\sin(t)}{t} dt = \sum_{k=0}^{\infty} (-1)^k L_k(x) < \infty. \quad (20.732)$$

*Démonstration.* Nous fixons (provisoirement)  $b$  et nous découpons l'intervalle d'intégration comme

$$[0, b] = \bigcup_{k=1}^N [k\pi, (k+1)\pi] \cup [(N+1)\pi, b] \quad (20.733)$$

où  $N$  est une fonction de  $b$  ; quelque chose comme  $N(b)$  est le plus grand entier tel que  $(N(b)+1)\pi \leq b$ . Sur chacun des intervalles nous avons  $\sin(t) = (-1)^k |\sin(t)|$ . Nous avons donc

$$\begin{aligned} \int_0^b (1 - e^{-tx}) \frac{\sin(t)}{t} dt &= \sum_{k=0}^{N(b)} (-1)^k \int_{k\pi}^{(k+1)\pi} (1 - e^{-tx}) \frac{|\sin(t)|}{t} dt \\ &\quad + \int_{(N+1)\pi}^b (1 - e^{-tx}) \frac{\sin(t)}{t} dt \end{aligned} \quad (20.734)$$

Le premier terme est  $\sum_{k=0}^{N(b)} (-1)^k L_k(x)$ , dont nous savons que la limite  $b \rightarrow \infty$  existe parce que  $L_k(x)$  vérifie le critère des séries alternées (lemme 20.218). En ce qui concerne le second terme,

$$\left| \int_{(N+1)\pi}^b (1 - e^{-tx}) \frac{\sin(t)}{t} dt \right| < \frac{b - (N+1)\pi}{(N+1)\pi} < \frac{1}{N(b) + 1}. \quad (20.735)$$

La dernière inégalité est le fait que  $N(b)$  est choisi pour avoir  $b - (N(b) + 1)\pi < \pi$ .

Les deux termes de (20.734) ont donc une limite lorsque  $b \rightarrow \infty$ . Nous pouvons donc passer à la limite en sommant les deux limites :

$$\int_0^\infty (1 - e^{-tx}) \frac{\sin(t)}{t} dt = \sum_{k=0}^\infty (-1)^k L_k(x). \quad (20.736)$$

Cette égalité est valable pour chaque  $x > 0$ .

Le fait que la limite soit finie est dans le critère des séries alternées. Pour chaque  $x$ , la suite  $L_k(x)$  vérifie ce critère par le lemme 20.218.  $\square$

LEM00NZVS00DbZCZx

### Lemme 20.220.

Nous avons

$$\lim_{x \rightarrow 0^+} \int_0^\infty (1 - e^{-tx}) \frac{\sin(t)}{t} dt = 0. \quad (20.737)$$

*Démonstration.* La définition de l'intégrale ainsi que le lemme 20.219 nous ont déjà donné

$$\int_0^\infty (1 - e^{-tx}) \frac{\sin(t)}{t} dt = \lim_{b \rightarrow \infty} \int_0^b (1 - e^{-tx}) \frac{\sin(t)}{t} dt = \sum_{k=0}^\infty (-1)^k L_k(x) \quad (20.738)$$

ainsi que l'assurance que le tout est un nombre réel fini<sup>48</sup>.

#### (1) Majoration pour la série alternée

Nous majorons un peu. Pour  $x > 0$  et  $N \in \mathbb{N}$  nous avons |

SUBEQSQ00LIGN00NAzpmi

$$\left| \sum_{k=0}^N (-1)^k L_k(x) \right| \leq \sum_{k=0}^N |L_k(x)| \quad (20.739a)$$

$$= \sum_{k=0}^N \left| \int_{k\pi}^{(k+1)\pi} (1 - e^{-tx}) \frac{\sin(t)}{t} dt \right| \quad (20.739b)$$

EQ00IIVV00LyHFhS

$$\leq \sum_{k=0}^N \int_{k\pi}^{(N+1)\pi} (1 - e^{-tx}) \frac{|\sin(t)|}{t} dt \quad (20.739c)$$

$$= \int_0^{(N+1)\pi} (1 - e^{-tx}) \frac{|\sin(t)|}{t} dt \quad (20.739d)$$

$$\leq \int_0^{(N+1)\pi} tx \frac{|\sin(t)|}{t} dt \quad (20.739e)$$

SUBEQ00ZPVV00TBZzdl

$$\leq \int_0^{(N+1)\pi} x dt \quad (20.739f)$$

$$= x(N+1)\pi. \quad (20.739g)$$

Justifications :

- Pour (20.739b) c'est la définition (20.728).
- Pour (20.739e), c'est le fait que  $0 \leq 1 - e^{-u} \leq u$  pour tout  $u \geq 0$  ainsi que la sous-additivité de l'intégrale de la proposition 14.194.

48. De toutes façons, il n'existe pas de nombres réels infinis, mais vous voyez ce que je veux dire.

(2) **Majoration pour l'intégrale**

Nous fixons  $N \in \mathbb{N}$ , et nous avons : | SUBEQSooSBSJooMAkJPh

$$\left| \lim_{b \rightarrow \infty} \int_0^b (1 - e^{-tx}) \frac{\sin(t)}{t} dt \right| = \left| \sum_{k=0}^{\infty} (-1)^k L_k(x) \right| \quad (20.740a)$$

$$= \left| \sum_{k=0}^N (-1)^k L_k(x) + \sum_{k=N+1}^{\infty} (-1)^k L_k(x) \right| \quad (20.740b)$$

$$\leq \left| \sum_{k=0}^N (-1)^k L_k(x) \right| + \left| \sum_{k=N+1}^{\infty} (-1)^k L_k(x) \right| \quad (20.740c)$$

$$\leq \left| \sum_{k=0}^N (-1)^k L_k(x) \right| + L_{N+1}(x) \quad (20.740d) \quad \text{SUBEQooYFWDooYKhYtd}$$

$$\leq \left| \sum_{k=0}^N (-1)^k L_k(x) \right| + \frac{1}{N+1} \quad (20.740e) \quad \text{SUBEQooDDRGooNDfxgQ}$$

$$(20.740f)$$

Justifications :

- Pour (20.740d) c'est le reste du critère des séries alternées, théorème 10.163(3).
- Pour (20.740e) c'est la majoration (20.729) déjà faite.

- (3) **Les deux ensemble** Pour chaque  $N$  et pour chaque  $x > 0$  nous avons, en mettant (20.739) au bout de (20.740) :

$$\left| \int_0^{\infty} (1 - e^{-tx}) \frac{\sin(t)}{t} dt \right| \leq x(N+1)\pi + \frac{1}{N+1}. \quad (20.741)$$

En prenant la limite  $x \rightarrow 0$  nous trouvons

$$\lim_{x \rightarrow 0} \int_0^{\infty} (1 - e^{-tx}) \frac{\sin(t)}{t} dt \leq \frac{1}{N+1} \quad (20.742)$$

pour tout  $N$ . Donc cette limite est nulle :

$$\lim_{x \rightarrow 0} \int_0^{\infty} (1 - e^{-tx}) \frac{\sin(t)}{t} dt = 0. \quad (20.743)$$

□

Maintenant que nous avons fait plein de lemmes, nous pouvons énoncer notre résultat principal, et le démontrer facilement.

**Théorème 20.221** (Intégrale de Dirichlet[668]).

Nous avons

$$\lim_{b \rightarrow \infty} \int_0^b \frac{\sin(t)}{t} dt = \frac{\pi}{2}. \quad (20.744)$$

Nous avons écrit  $\lim_{b \rightarrow \infty} \int_0^b \frac{\sin(t)}{t} dt$  et non  $\int_0^{\infty} \frac{\sin(t)}{t} dt$  parce que cette dernière intégrale n'existe pas vraiment au sens de Lebesgue, voir le lemme 20.213. Dans la suite nous écrirons cependant  $\int_0^{\infty} \frac{\sin(t)}{t} dt$ , en gardant en tête que cela n'est défini que via la limite.

*Démonstration.* Nous nommons  $D$  la valeur que nous cherchons. Le lemme 20.214 nous assure que

$$D = \lim_{b \rightarrow \infty} \int_0^b \frac{\sin(t)}{t} dt < \infty. \quad (20.745)$$



Le lemme 20.217 nous donne, quant à lui,

$$\lim_{b \rightarrow \infty} \int_0^b e^{-tx} \frac{\sin(t)}{t} dt = -\arctan(x) + \frac{\pi}{2}. \quad (20.746)$$

Vu que les deux limites existent, on peut permuter somme et limite<sup>49</sup> :

$$\begin{aligned} D + \arctan(x) - \frac{\pi}{2} &= \lim_{b \rightarrow \infty} \left( \int_0^b \frac{\sin(t)}{t} + \int_0^b e^{-tx} \frac{\sin(t)}{t} dt \right) \\ &= \lim_{b \rightarrow \infty} \int_0^b (1 - e^{-xt}) \frac{\sin(t)}{t} dt. \end{aligned} \quad (20.747a)$$

(20.747b)

Pour (20.747b), nous avons des fonctions bornées sur un intervalle borné  $]0, b[$ , donc il n'y a pas de mal à sommer les intégrales.

Donc pour tout  $x > 0$ , nous avons

$$D + \arctan(x) - \frac{\pi}{2} = \int_0^\infty (1 - e^{-xt}) \frac{\sin(t)}{t} dt. \quad (20.748)$$

Nous passons à la limite  $x \rightarrow 0$  en utilisant le lemme 20.220 et le fait que  $\arctan(0) = 0$  (lemme 18.44) :

$$D - \frac{\pi}{2} = 0, \quad (20.749)$$

c'est-à-dire que résultat annoncé.  $\square$

---

49. C'est une phrase un peu grandiloquente pour dire que  $\lim_{b \rightarrow a} f(b) - \lim_{b \rightarrow a} g(b) = \lim_{b \rightarrow a} (f(b) - g(b))$ . Ici nous avons  $a = \infty$  et les fonctions  $f$  et  $g$  sont celles définies par les intégrales.

# Chapitre 21

## Arcs paramétrés

La structure de ce chapitre, comme beaucoup de choses dans le Frido, est fortement liée au choix de présenter toutes les matières dans l'ordre mathématiquement logique. Nous devons donc le placer après la trigonométrie ; les propriétés principales des fonctions trigonométriques étant dans la proposition 18.21, et c'est la proposition 18.59 qui nous permet de dire que  $(\cos(t), \sin(t))$  décrit le cercle.

Et enfin nous n'avons pas encore calculé la circonférence du cercle, et pour cause : nous n'avons pas encore donné de définition à la longueur d'un chemin dans  $\mathbb{R}^2$ . C'est pourquoi ce chapitre va aller droit à la longueur avant de donner des exemples.

### 21.1 Définitions

SecDeExCPar

#### Définition 21.1.

Un **arc paramétré** dans  $\mathbb{R}^p$  est un couple  $(I, \gamma)$  où  $I$  est un intervalle de  $\mathbb{R}$  et  $\gamma$  est une application continue de  $I$  dans  $\mathbb{R}^p$ . Nous disons que  $(I, \gamma)$  est un arc paramétré **compact** (ou un **chemin** dans  $\mathbb{R}^p$ ) lorsque  $I$  est compact dans  $\mathbb{R}$ .

L'intervalle  $I$  d'un arc paramétré compact est toujours de la forme  $[a, b]$ , étant donné que tous les intervalles compacts de  $\mathbb{R}$  sont de cette forme. Un **sous arc** de  $(I, \gamma)$  est un arc de la forme  $(I_0, \gamma)$  avec  $I_0 \subset I$ .

La définition 20.75 a déjà dit qu'un chemin dans  $\mathbb{R}^n$  est une application continue  $\sigma : [a, b] \rightarrow \mathbb{R}^n$ .

Si le chemin  $\sigma$  est dérivable, la fonction  $\sigma'$  est la **vitesse** du chemin  $\sigma$ . Si le chemin est deux fois dérivable, la fonction  $\sigma''$  l'**accélération**.

L'image de  $\sigma$  dans  $\mathbb{R}^3$  est ce qu'on appellera la **courbe**.

### 21.2 Longueur d'arc

SecLongArc

Nous voulons définir et étudier la notion de **longueur** d'un arc paramétré. Pour cela, le plus raisonnable est d'approcher l'arc par des petits segments de droites (dont les longueurs sont évidentes), et d'extraire la « meilleure » approximation.

Une des notions clefs pour la suite est celle de subdivision d'intervalles. Cette notion sera encore utilisée par la suite à propos des intégrales.

DefSubdivisionIntervalle

#### Définition 21.2.

Si  $I$  est un intervalle d'extrêmes  $a$  et  $b$  avec  $a < b$ , nous appelons **subdivision finie** de  $I$  un choix de nombres  $t_i$  tels que

$$a = t_0 < t_1 < \dots < t_n = b. \quad (21.1)$$

Nous disons qu'une subdivision  $\sigma'$  est **plus fine** que la subdivision  $\sigma$  si l'ensemble des points de  $\sigma$  est inclus dans celui des points de  $\sigma'$ . Dans ce cas, la subdivision  $\sigma'$  est un **raffinement** de  $\sigma$ . Nous désignons par  $\mathcal{S}(I)$  l'ensemble des subdivisions finies de l'intervalle  $I$ .

Dans la suite, toutes les subdivisions que nous considérons seront des subdivisions finies. Aussi nous parlerons simplement de *subdivisions* sans préciser. Nous allons souvent noter  $\sigma = (t_i)_{i=1}^n$  pour désigner la subdivision formée par les nombres  $t_i$ . Il faut garder en tête que dans une subdivision, les nombres *sont ordonnés*.

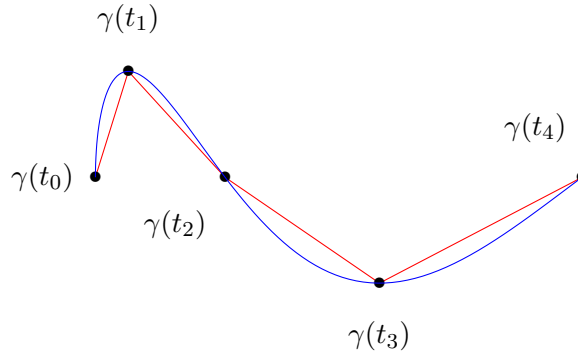


FIGURE 21.1: La longueur d'un découpage. La somme des longueurs des segments droits est facile à calculer.

LabelFigCourbeRectifiable

DEFooDNZWooXmxhsU

### Définition 21.3.

Soit un arc paramétré compact  $(I, \gamma)$  et une subdivision  $\sigma = (t_i)_{i=0}^n$  de  $I = [a, b]$ . À partir de  $\gamma$  et du découpage  $\sigma$  nous définissons le nombre (voir figure 21.1)

$$l_\sigma(\gamma) = \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\|. \quad (21.2)$$

Eqlsigmagamasss

On appelle **longueur** de l'arc  $\gamma$  le nombre

$$l(\gamma) = \sup_{\sigma} l_\sigma(\gamma) \in [0, \infty]. \quad (21.3)$$

Nous disons que  $\gamma$  est **rectifiable** lorsque  $l(\gamma) < \infty$ .

Lorsque nous voulons spécifier sur quel intervalle nous considérons l'arc, nous noterons  $l(I, \gamma)$  au lieu de  $l(\gamma)$  pour être plus précis.

Par l'inégalité triangulaire, si  $\sigma_1$  est plus fine que  $\sigma$ , nous avons

$$l_\sigma(\gamma) \leq l_{\sigma_1}(\gamma), \quad (21.4)$$

Comme cela peut être vu sur la figure 21.2.

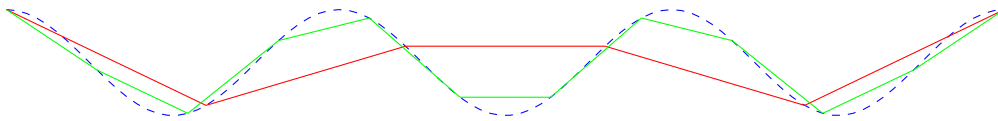


FIGURE 21.2: Il est visible que la longueur donnée par l'approximation par des petits segments (verts) est plus longue et plus précise que celle donnée par les longs segments (rouge).

LabelFigArcLongueurFinesse

PROPoCXLyooRpKDMs

### Proposition 21.4.

Si  $P$  et  $Q$  sont des points de  $\mathbb{R}^2$ , alors le segment de droite joignant  $P$  à  $Q$  est le plus court des arcs paramétrés passant par  $P$  et  $Q$ .

*Démonstration.* Si  $\gamma$  est un arc paramétré joignant  $P$  et  $Q$ , la longueur de  $\gamma$  est donné par un supremum dont un des éléments est la longueur du segment de droite.  $\square$

Dans la vie réelle, il est souvent difficile et peu pratique de calculer le supremum « à la main ». C'est pourquoi nous allons travailler à exprimer la longueur d'un arc à l'aide d'une intégrale (théorème 21.9).

**Lemme 21.5.**

Nous avons  $l(\gamma) = 0$  si et seulement si  $\gamma(t)$  est un vecteur constant.

*Démonstration.* Si l'application  $\gamma(t)$  est constante, le résultat est évident. Supposons maintenant que  $\gamma$  ne soit pas constante. Cela signifie qu'il existe  $t_1$  et  $t_2$  dans  $I$  tels que  $\gamma(t_1) \neq \gamma(t_2)$ . Dans ce cas, si nous prenons le découpage  $\sigma = \{a, t_1, t_2, b\}$ , la somme (21.2) contient au moins le terme non nul  $\|\gamma(t_2) - \gamma(t_1)\|$ , et donc  $l_\sigma(\gamma) > 0$ . Par définition du supremum, nous avons alors  $l(\gamma) \geq l_\sigma(\gamma) > 0$ .  $\square$

Problemautredcop

**Proposition 21.6.**

Soit  $(I, \gamma)$  un arc paramétré compact.

- (1) Si  $\gamma' = (I', \gamma)$  avec  $I' \subset I$ , alors  $l(\gamma') \leq l(\gamma)$ .
- (2) Soit  $c \in [a, b]$ , et considérons les arcs  $\gamma_1 = ([a, c], \gamma)$  et  $\gamma_2 = ([c, b], \gamma)$ . Alors

$$l(\gamma) = l(\gamma_1) + l(\gamma_2). \quad (21.5)$$

En particulier,  $\gamma$  est rectifiable si et seulement si  $\gamma_1$  et  $\gamma_2$  le sont.

*Démonstration.* (1) Nous notons  $I = [a, b]$  et  $I' = [a', b']$ . Étant donné que  $I' \subset I$ , nous avons

$$a \leq a' < b' \leq b. \quad (21.6)$$

Pour chaque subdivision  $\sigma_0 : a' = t_0 < t_1 < \dots < t_n = b'$  de  $I'$ , nous pouvons construire une subdivision de  $I$  en « ajoutant » les points  $a$  et  $b$ , c'est-à-dire

$$\sigma : a \leq t_0 < \dots < t_n \leq b. \quad (21.7)$$

Si nous calculons  $l_\sigma(\gamma)$ , nous avons tous les termes qui arrivent dans  $l_{\sigma_0}(\gamma')$  plus le premier et dernier terme :  $\|\gamma(t_0) - \gamma(a)\|$  et  $\|\gamma(b) - \gamma(t_n)\|$ . Nous avons donc

$$l_{\sigma_0}(\gamma') \leq l_\sigma(\gamma) \leq \sup_{\sigma} l_\sigma(\gamma) = l(\gamma). \quad (21.8)$$

Étant donné que pour toute subdivision  $\sigma_0$  nous avons  $l_{\sigma_0}(\gamma') \leq l(\gamma)$ , en prenant le supremum sur les subdivisions  $\sigma_0$  de  $I'$ , nous avons comme annoncé

$$l(\gamma') \leq l(\gamma). \quad (21.9)$$

- (2) Soit  $\sigma = \{t_i\}$  une subdivision de  $[a, b]$ . Nous considérons les subdivisions  $\sigma_1$  et  $\sigma_2$  définies comme suit :

$$\begin{aligned} \sigma_1 &: \{t_i \text{ tel que } t_i < c\} \cup \{c\}, \\ \sigma_2 &: \{t_i \text{ tel que } t_i > c\} \cup \{c\}. \end{aligned} \quad (21.10)$$

L'inégalité triangulaire implique que

$$l_\sigma(\gamma) \leq l_{\sigma \cup \{c\}}(\gamma) = l_{\sigma_1}(\gamma_1) + l_{\sigma_2}(\gamma_2) \leq l(\gamma_1) + l(\gamma_2). \quad (21.11)$$

Nous avons donc

$$l(\gamma) \leq l(\gamma_1) + l(\gamma_2). \quad (21.12)$$

EqIneqIglglglud

Nous prouvons maintenant l'inégalité inverse. Soit  $\varepsilon > 0$ . Étant donné que  $l(\gamma_1)$  est le supremum des quantités  $l_{\sigma_1}(\gamma_1)$  lorsque  $\sigma_1$  parcourt toutes les subdivisions possibles, il existe une partition  $\sigma_1^\varepsilon$  telle que (idem pour  $\gamma_2$ )

$$\begin{aligned} l_{\sigma_1^\varepsilon}(\gamma_1) + \frac{\varepsilon}{2} &> l(\gamma_1), \\ l_{\sigma_2^\varepsilon}(\gamma_2) + \frac{\varepsilon}{2} &> l(\gamma_2), \end{aligned} \quad (21.13)$$

EqAllsigmaepsgammaaufd

où  $\sigma_1^\varepsilon$  est une subdivision de  $[a, c]$  et  $\sigma_2^\varepsilon$  en est une de  $[c, b]$ . En faisant la somme des deux équations (21.13), nous trouvons

$$l(\gamma_1) + l(\gamma_2) < l_{\sigma_1^\varepsilon}(\gamma_1) + l_{\sigma_2^\varepsilon}(\gamma_2) + \varepsilon = l_{\sigma_1^\varepsilon \cup \sigma_2^\varepsilon}(\gamma) \leq l(\gamma) + \varepsilon. \quad (21.14)$$

L'inégalité  $l(\gamma_1) + l(\gamma_2) < l(\gamma) + \varepsilon$  étant valable pour tout  $\varepsilon$ , nous avons

$$l(\gamma_1) + l(\gamma_2) \leq l(\gamma). \quad (21.15)$$

Cette inégalité, combinée avec l'inégalité (21.12), donne bien  $l(\gamma) = l(\gamma_1) + l(\gamma_2)$ . □

## 21.3 Abscisse curviligne

### Définition 21.7.

Soit  $(I, \gamma)$  un arc rectifiable compact avec  $I = [a, b]$ . L'application

$$\begin{aligned} \varphi: [a, b] &\rightarrow \mathbb{R}^+ \\ t &\mapsto l([a, t], \gamma) \end{aligned} \quad (21.16)$$

est la **longueur d'arc** de  $\gamma$ .

Cette fonction nous permet de calculer la distance (suivant la courbe) entre deux points arbitraires parce que si  $a \leq t < u \leq b$ , nous avons

$$l([t, u], \gamma) = \varphi(u) - \varphi(t). \quad (21.17)$$

En effet,

$$\varphi(u) - \varphi(t) = l([a, u], \gamma) - l([a, t], \gamma), \quad (21.18)$$

mais en utilisant la proposition 21.6, nous avons

$$l([a, u], \gamma) = l([a, t], \gamma) + l([t, u], \gamma). \quad (21.19)$$

### Proposition 21.8.

La longueur d'arc d'un arc rectifiable compact est une fonction continue et croissante.

*Démonstration.* Soit  $(I, \gamma)$  un arc paramétré rectifiable compact avec  $I = [a, b]$ . Afin de montrer que  $\varphi$  est croissante, prenons  $t \in I$  ainsi que  $h > 0$  et montrons que  $\varphi(t + h) \geq \varphi(t)$ . La proposition 21.6 implique que

$$l([a, t + h], \gamma) = l([a, t], \gamma) + l([t, t + h], \gamma), \quad (21.20)$$

c'est-à-dire

$$\varphi(t + h) = \varphi(t) + l([t, t + h], \gamma) \geq \varphi(t). \quad (21.21)$$

Pour la continuité, soit  $t$  fixé dans  $[a, b]$  et  $\varepsilon > 0$ . Il nous faut démontrer qu'il existe  $\eta > 0$  tel que si  $s$  est dans  $[0, \eta]$  alors

$$|\varphi(t + s) - \varphi(t)| \leq \varepsilon, \quad \forall t \in [a, b].$$

Étant donné que  $l([t, b], \gamma)$  est le supremum des  $l_\sigma([t, b], \gamma)$ , il existe une subdivision  $\sigma$  donnée par les points  $t, t_1, \dots, t_{n-1}, b$  telle que

$$l_\sigma([t, b], \gamma) > l([t, b], \gamma) - \frac{\varepsilon}{2} = \varphi(b) - \varphi(t) - \frac{\varepsilon}{2}. \quad (21.22)$$

La continuité de  $\gamma$  implique qu'il existe un  $\eta$  tel que

$$s \in [0, \eta] \Rightarrow \|\gamma(t + s) - \gamma(t)\| < \frac{\varepsilon}{2} \quad (21.23)$$

Quitte à prendre  $\eta$  encore plus petit, nous supposons que  $t + \eta < t_1$ . Soit  $s \in [0, \eta]$  et considérons la subdivision de  $[t, b]$  donnée par  $\sigma' = \sigma \cup \{t + s\}$ . Étant donné que  $\sigma'$  est plus fine que  $\sigma$ , le nombre  $l_\sigma([t, b], \gamma)$  est inférieur ou égal à  $l_{\sigma'}([t, b], \gamma)$ . Nous avons donc les inégalités

$$\begin{aligned} \varphi(b) - \varphi(t) - \frac{\varepsilon}{2} &\leq l_\sigma([t, b], \gamma) \\ &\leq l_{\sigma'}([t, b], \gamma) \\ &= \|\gamma(t + s) - \gamma(t)\| + l_{\sigma' \setminus \{t\}}([t + s, b], \gamma) \\ &\leq \|\gamma(t + s) - \gamma(t)\| + \varphi(b) - \varphi(t + s) \\ &\leq \frac{\varepsilon}{2} + \varphi(b) - \varphi(t + s). \end{aligned} \quad (21.24)$$

Au final, nous avons trouvé que

$$\varphi(t + s) - \varphi(t) \leq \varepsilon, \quad (21.25)$$

ce qui prouve que  $\varphi$  est continue au point  $t$ .  $\square$

En guise de paramètre sur un arc, nous pouvons utiliser la longueur d'arc elle-même. En effet si  $(I, \gamma)$  est un arc de longueur  $l$ , nous pouvons donner le même arc avec le couple  $([0, l], g)$  où  $g$  est la fonction qui au réel  $s$  fait correspondre l'élément  $\gamma(\varphi^{-1}(s))$  de  $\mathbb{R}^n$ . Dire

$$P = (\gamma \circ \varphi^{-1})(s) \quad (21.26)$$

revient à dire que le point  $P$  est le point sur la courbe sur lequel on tombe après avoir marché une distance  $s$  sur la courbe.

Nous allons revenir sur ce « changement de paramètre » plus tard, en particulier dans la section 21.7.

### 21.3.1 Formule intégrale de la longueur

Nous pouvons voir un chemin  $\gamma$  comme étant la trajectoire d'une particule en fonction du temps. Sa vitesse à l'instant  $t$  est le vecteur  $\gamma'(t)$ , tandis que sa vitesse *scalair*e est le nombre  $\|\gamma'(t)\|$ . Une question naturelle est de savoir quelle est la longueur de la trajectoire parcourue entre  $t = a$  et  $t = b$ .

Si nous prenons un petit intervalle de temps  $dt$ , nous pouvons supposer que le mobile avance à la vitesse constante  $\|\gamma'(t)\|$ . Cela ferait un trajet parcouru de longueur  $\|\gamma'(t)\|dt$ . Nous nous attendons donc à une formule de la forme suivante pour la longueur de  $\gamma$  :

$$l(\gamma) = \int_a^b \|\gamma'(t)\| dt. \quad \text{EqDefLongueurChemin} \quad (21.27)$$

Plus explicitement, si  $\gamma(t) = (x(t), y(t), z(t))$ , alors nous aurions la formule

$$l(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt. \quad (21.28)$$

#### Théorème 21.9.

Soit  $(I, \gamma)$  un arc paramétré compact de classe  $\mathcal{C}^1$ . Alors  $\gamma$  est rectifiable et

$$l(\gamma) = \int_a^b \|\gamma'(t)\| dt = \int_\gamma 1, \quad \text{EqLongGammaInt} \quad (21.29)$$

où  $I = [a, b]$ .

*Démonstration.* L'égalité avec l'intégrale le long de  $\gamma$  de la fonction 1 est simplement la définition 20.40 de l'intégrale curviligne.

Si  $\sigma = \{t_i\}$  est une subdivision de l'intervalle  $[a, b]$ , alors

$$\begin{aligned} l_\sigma(\gamma) &= \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\| \\ &= \sum_{i=1}^n \left\| \int_{t_{i-1}}^{t_i} \gamma'(t) dt \right\| \\ &\leq \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \|\gamma'(t)\| dt \\ &= \int_a^b \|\gamma'(t)\| dt. \end{aligned} \tag{21.30}$$

Cela prouve déjà que

$$l(\gamma) = \sup_\sigma l_\sigma(\gamma) \leq \int_a^b \|\gamma'(t)\| dt. \tag{21.31}$$

Nous devons maintenant prouver l'inégalité inverse.

Notons  $\varphi$  l'abscisse curviligne  $\varphi(t) = l([a, t], \gamma)$ . Cette dernière vérifie

$$\varphi(t+h) - \varphi(t) = l([t, t+h], \gamma) \geq \|\gamma(t+h) - \gamma(t)\|, \tag{21.32}$$

et en particulier

$$\left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| \leq \frac{\varphi(t+h) - \varphi(t)}{h}. \tag{21.33}$$

D'autre part, en utilisant (21.31) sur le segment  $[t, t+h]$ , nous avons

$$\varphi(t+h) - \varphi(t) = l([t, t+h], \gamma) \leq \int_t^{t+h} \|\gamma'(u)\| du. \tag{21.34}$$

Cela nous permet de continuer l'inéquation (21.33) en

$$\left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| \leq \frac{\varphi(t+h) - \varphi(t)}{h} \leq \frac{1}{h} \int_t^{t+h} \|\gamma'(u)\| du. \tag{21.35}$$

Prenons la limite  $h \rightarrow 0$ . À gauche nous reconnaissons la formule de la dérivée, et nous obtenons  $\|\gamma'(t)\|$ ; au centre nous avons  $\varphi'(t)$  et à droite, si  $n(u)$  représente une primitive de la fonction  $u \mapsto \|\gamma'(u)\|$ ,

$$\lim_{h \rightarrow 0} \frac{n(t+h) - n(t)}{h} = n'(t) = \|\gamma'(t)\|. \tag{21.36}$$

Au final,

$$\|\gamma'(t)\| \leq \varphi'(t) \leq \|\gamma'(t)\|, \tag{21.37}$$

c'est-à-dire  $\varphi'(t) = \|\gamma'(t)\|$  et donc par le théorème fondamental du calcul intégral 14.288,

$$\varphi(t) - \varphi(a) = \int_a^t \|\gamma'(u)\| du. \tag{21.38}$$

Par construction de la longueur d'arc,  $\varphi(a) = 0$  et en posant  $t = b$  nous obtenons la relation recherchée :

$$l(\gamma) = \varphi(b) = \int_a^b \|\gamma'(u)\| du. \tag{21.39}$$

□

RemLongIntUn

### Remarque 21.10.

Cela est cohérent avec 20.186, mais il faut garder en tête que  $l(\gamma)$  n'est pas la mesure de Lebesgue de l'image de  $\gamma$  dans  $\mathbb{R}^2$ . Cette dernière est nulle.

**Exemple 21.11.**

Soient donc  $a$  et  $b$  deux points de  $\mathbb{R}^m$ , et  $\gamma$  la droite joignant  $a$  à  $b$ , c'est-à-dire

$$\gamma(t) = (1-t)a + tb \quad (21.40)$$

avec  $t \in [0, 1]$ . Le théorème 21.9 nous enseigne que la longueur de ce chemin est

$$l([0, 1], \gamma) = \int_0^1 \|\gamma'(t)\| dt = \int_0^1 \| -a + b \| = \|b - a\|, \quad (21.41)$$

qui est bien la distance entre  $a$  et  $b$ . △

**Exemple 21.12** (Circonférence du cercle).

Nous savons que l'image de

$$\begin{aligned} \gamma: [0, 2\pi[ &\rightarrow \mathbb{R}^2 \\ t &\mapsto (R \cos(t), R \sin(t)) \end{aligned} \quad (21.42)$$

est le cercle de centre  $(0, 0)$  et de rayon  $R > 0$ . Et de plus cet arc est de classe  $C^1$  (et même  $C^\infty$ ) par la proposition 18.1. La longueur sera, d'après la formule (21.9)

$$l_\gamma = \int_0^{2\pi} \|\gamma'(t)\| dt = 2\pi R \quad (21.43)$$

grâce à la formule  $\sin^2 + \cos^2 = 1$  du lemme 18.4.

Mais tout cela n'est pas satisfaisant parce que nous n'avons pas encore de valeur numérique de  $\pi$ .

Il y a une autre façon de faire en considérant le quart de cercle dont la longueur en fonction de  $\pi$  est vite calculée par

$$l_\gamma = \int_0^{\pi/2} \|\gamma'(t)\| dt = \frac{\pi R}{2}. \quad (21.44)$$

Cette même longueur est calculée en termes de fonctions plus courantes avec le chemin

$$\begin{aligned} \sigma: ]0, 1[ &\rightarrow \mathbb{R}^2 \\ t &\mapsto \left( \frac{t}{\sqrt{R^2 - t^2}} \right) \end{aligned} \quad (21.45)$$

La longueur s'exprime avec

$$l_\sigma = \int_0^1 \sqrt{\frac{R^2}{R^2 + t^2}} dt. \quad (21.46)$$

Notons que le changement de variables  $t = R \sin(u)$  permet de retrouver l'expression  $l_\sigma = \pi R/2$ .

Pour avoir une approximation de  $\pi$ , il est loisible de calculer une approximation numérique de l'intégrale (21.46) (avec  $R = 1$ ) et de l'égaliser à  $\pi/2$ . △

**Proposition 21.13.**

Soit le cercle  $S = \{x \in \mathbb{R}^2 \text{ tel que } \|x\| = 1\}$ . Nous considérons la bijection<sup>1</sup>  $\varphi: [0, 2\pi[ \rightarrow S$ . Soient deux points  $a, b \in S$  ainsi que  $\theta_a = \varphi^{-1}(a)$ ,  $\theta_b = \varphi^{-1}(b)$ . Nous supposons que  $\theta_a \leq \theta_b$  et que  $\widehat{a(0,0)b} = \theta$ <sup>2</sup>.

Nous considérons le chemin

$$\begin{aligned} \gamma: [\theta_a, \theta_b] &\rightarrow S \\ t &\mapsto (R \cos(t), R \sin(t)). \end{aligned} \quad (21.47)$$

Alors  $l(\gamma) = R\theta$ <sup>3</sup>.

- 
1. Corolaire 18.60.
  2. Angle, définition 18.156
  3. Longueur d'arc, définition 21.3.



*Démonstration.* Vu que  $\|a\| = \|b\| = R$ , la définition de l'angle  $\widehat{aob}$  dit que

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \frac{a}{R} = \frac{b}{R}. \quad (21.48)$$

En substituant  $a = R(\cos(\theta_a), \sin(\theta_a))$  et  $b = R(\cos(\theta_b), \sin(\theta_b))$  et en faisant une petite multiplication matrice par vecteur :

$$\begin{pmatrix} \cos(\theta) \cos(\theta_a) - \sin(\theta) \sin(\theta_a) \\ \sin(\theta) \cos(\theta_a) + \cos(\theta) \sin(\theta_a) \end{pmatrix} = \begin{pmatrix} \cos(\theta_b) \\ \sin(\theta_b) \end{pmatrix} \quad (21.49)$$

En utilisant les formules de trigonométrie du lemme 18.14, il vient

$$\begin{cases} \cos(\theta + \theta_a) = \cos(\theta_b) \\ \sin(\theta + \theta_a) = \sin(\theta_b). \end{cases} \quad \begin{matrix} (21.50a) \\ (21.50b) \end{matrix}$$

La proposition 18.24 nous enseigne alors que  $\theta_b = \theta + \theta_a + 2k\pi$  pour un certain  $k \in \mathbb{Z}$ . Vu que  $\theta_a$  et  $\theta_b$  sont entre 0 et  $2\pi$ , il est obligatoire que  $k = 0$  et donc que  $\theta_b = \theta_a + \theta$ .

Enfin nous calculons la longueur de  $\gamma$  avec le théorème 21.9 qui permet de réduire la longueur à une intégrale :  $l(\gamma) = \int_{\theta_a}^{\theta_a+\theta} \|\gamma'(t)\| dt$ .

Dans notre cas,  $\gamma'(t) = R(-\sin(t), \cos(t))$  et donc  $\|\gamma'(t)\| = R$ . Donc

$$l(\gamma) = \int_{\theta_a}^{\theta_a+\theta} \|\gamma'(t)\| dt = R\theta, \quad (21.51)$$

comme annoncé. □

### Exemple 21.14.

Considérons l'arc de cercle de rayon  $R$  interceptée par l'angle  $\theta$  présenté sur la figure 21.3.

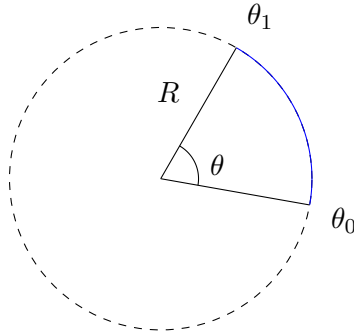


FIGURE 21.3: Quelle est la longueur de la partie bleue de ce cercle de rayon  $R$ ?

Par définition, cette longueur sera

$$\int_{\theta_0}^{\theta_1} \sqrt{R^2 \sin^2(t) + R^2 \cos^2(t)} dt = R(\theta_1 - \theta_0). \quad (21.52)$$

Le radian comme unité de mesure d'angle est donc l'unité parfaite : elle est la longueur d'arc interceptée (si le rayon est  $R = 1$ ). △

Une conséquence à peine indirecte de ce que nous venons de voir à propos de longueur d'arc de cercle est la proposition suivante<sup>4</sup>.

### Proposition 21.15.

Pour tout  $x, y \in \mathbb{R}$ , nous avons

$$|e^{ix} - e^{iy}| \leq |x - y|. \quad (21.53)$$

4. À mon avis il y a moyen de prouver ça avec un développement limité, mais je ne sais pas trop comment majorer l'erreur sans accepter que  $x$  soit arbitrairement proche de  $y$ . Si vous savez comment faire, écrivez-moi.

*Démonstration.* Évacuons tout de suite la différence entre  $\mathbb{R}^2$  et  $\mathbb{C}$  : ils sont isométriques. Si vous n'êtes pas convaincu que tout se passe bien, vous pouvez récrire toute la démonstration en écrivant systématiquement  $(\cos(x), \sin(x))$  au lieu de  $e^{ix}$ . Cela serait au passage un bon exercice pour voir que les formules de dérivation fonctionnent bien.

Nous considérons les points  $e^{ix}$  et  $e^{iy}$  dans  $\mathbb{C}$  et deux chemins différents les joignant. Le premier est le segment de droite

$$\begin{aligned}\sigma_1 : [0, 1] &\rightarrow \mathbb{C} \\ t &\mapsto te^{ix} + (1-t)e^{iy}.\end{aligned}\tag{21.54}$$

Le second est l'arc de cercle

$$\begin{aligned}\sigma_2 : [0, 1] &\rightarrow \mathbb{C} \\ t &\mapsto e^{i(tx+(1-t)y)}.\end{aligned}\tag{21.55}$$

Nous avons  $\sigma_1'(t) = e^{ix} - e^{iy}$  qui ne dépend pas de  $t$ , et donc la longueur est facile à calculer à partir de la formule intégrale du théorème 21.9 :

$$l(\sigma_1) = \int_0^1 |\sigma_1'(t)| = |e^{ix} - e^{iy}|.\tag{21.56}$$

En ce qui concerne le second chemin,

$$\sigma_2'(t) = (x - y)e^{i(tx+(1-t)y)}.\tag{21.57}$$

Nous avons <sup>5</sup>  $|\sigma_2'(t)| = |x - y|$  qui ne dépend pas non plus de  $t$ . Donc

$$l(\sigma_2) = |x - y|.\tag{21.58}$$

Étant donné la proposition 21.4 qui dit que le chemin le plus court est le segment de droite,

$$l(\sigma_1) < l(\sigma_2)\tag{21.59}$$

et donc le résultat annoncé.  $\square$

### 21.16.

Si on veut savoir la longueur d'une courbe donnée sous la forme d'une fonction  $y = y(x)$ , un chemin qui trace la courbe est évidemment donné par

$$\gamma(t) = (t, y(t)),\tag{21.60}$$

et le vecteur tangent au chemin est  $\gamma'(t) = (1, y'(t))$ . Donc

$$\|\gamma'(t)\| = \sqrt{1 + y'(t)^2},\tag{21.61}$$

et

$$L = \int_a^b \sqrt{1 + y'(t)^2}.\tag{21.62}$$

*EqLongFonction*

### Exemple 21.17.

La longueur de l'hélice

$$\sigma(t) = \begin{pmatrix} \cos(2t) \\ \sin(2t) \\ \sqrt{5}t \end{pmatrix}\tag{21.63}$$

pour  $t \in [0, 4\pi]$  est donnée par

$$l(\sigma) = \int_0^{4\pi} \sqrt{4\sin^2(2t) + 4\cos^2(2t) + 5} dt = \int_0^{4\pi} \sqrt{9} = 12\pi.\tag{21.64}$$

$\triangle$

---

5. Si vous voulez citer des résultats, lemme 18.11 et proposition 9.104.

**Définition 21.18.**

Soit  $\sigma_1: [a, b] \rightarrow \mathbb{R}^3$ , un chemin et  $\sigma_2: [c, d] \rightarrow \mathbb{R}^3$ , un autre chemin. On dit que ces chemins sont **équivalents** si il existe une fonction  $\varphi: [a, b] \rightarrow [c, d]$  strictement croissante telle que  $\sigma_1(t) = \sigma_2(\varphi(t))$ .

Deux chemins équivalents parcourent la même courbe dans le même sens. Ils ne le parcourent toutefois pas à la même vitesse. On dit que les chemins sont **opposés** si la fonction  $\varphi$  de la définition est strictement décroissante. Dans ce cas, ils ont la même image, mais parcourue dans le sens opposés. Nous disons que deux chemins équivalents sont un **changement de paramétrage** pour la même courbe.

Dans le cas d'un paramétrage équivalente, nous avons  $\varphi(a) = c$  et  $\varphi(b) = d$ . Les points de départ et d'arrivée des deux paramètres coïncident. Dans le cas d'un paramètre qui va dans le sens opposé par contre nous avons automatiquement  $\varphi(a) = d$  et  $\varphi(b) = c$ .

**Proposition 21.19.**

La longueur d'une courbe ne dépend pas du paramètre (équivalent ou opposé) choisi.

*Démonstration.* Soient  $\sigma_1: [a, b] \rightarrow \mathbb{R}^3$  et  $\sigma_2: [c, d] \rightarrow \mathbb{R}^3$  tels que

$$\sigma_1(t) = \sigma_2(\varphi(t)) \quad \text{EqChmsigmaundeuuxxp} \quad (21.65)$$

où  $\varphi: [a, b] \rightarrow [c, d]$  est une bijection strictement monotone. Par définition on a

$$l(\sigma_1) = \int_a^b \|\sigma_1'(t)\| dt. \quad (21.66)$$

Nous pouvons exprimer la dérivée de  $\sigma_1$  en termes de celle de  $\sigma_2$  en dérivant la relation (21.65) :

$$\sigma_1'(t) = \varphi'(t) \sigma_2'(\varphi(t)). \quad (21.67)$$

En ce qui concerne la norme,

$$\|\sigma_1'(t)\| = |\varphi'(t)| \|\sigma_2'(\varphi(t))\|. \quad (21.68)$$

Notez dans cette relation que  $\varphi'(t)$  est un nombre (et non un vecteur). Étant donné que nous avons supposé que  $\varphi$  était monotone, soit elle est monotone croissante et  $\|\varphi'(t)\| = \varphi'(t)$  pour tout  $t$ , soit elle est monotone décroissante et  $\|\varphi'(t)\| = -\varphi'(t)$  pour tout  $t$ .

Considérons d'abord le premier cas, c'est-à-dire  $\|\varphi'(t)\| = \varphi'(t)$ . Nous posons  $s = \varphi(t)$ ,  $ds = \varphi'(t)dt$ . En remplaçant cela dans la formule de la longueur est

$$\begin{aligned} l(\sigma_1) &= \int_a^b \varphi'(t) \|\sigma_2'(\varphi(t))\| dt \\ &= \int_{\varphi(a)}^{\varphi(b)} \|\sigma_2'(s)\| ds \\ &= \int_c^d \|\sigma_2'(s)\| ds \\ &= l(\sigma_2). \end{aligned} \quad (21.69)$$

Si nous considérons maintenant un paramétrage strictement décroissante. Dans ce cas,  $\varphi'(t) \leq 0$  et  $\|\varphi'(t)\| = -\varphi'(t)$ . Nous posons encore une fois  $s = \varphi(t)$ ,  $ds = \varphi'(t)dt$ . Ici il ne faut pas oublier

que  $\varphi(a) = d$  et  $\varphi(b) = c$ . Le calcul est à part cela le même en faisant attention au signe :

$$\begin{aligned}
 l(\sigma_1) &= \int_a^b \varphi'(t) \|\sigma'_2(\varphi(t))\| dt \\
 &= - \int_{\varphi(a)}^{\varphi(b)} \|\sigma'_2(s)\| ds \\
 &= - \int_d^c \|\sigma'_2(s)\| ds \\
 &= \int_c^d \|\sigma'_2(s)\| ds \\
 &= l(\sigma_2).
 \end{aligned} \tag{21.70}$$

Nous avons changé le signe en changeant l'ordre des bornes. □

## 21.4 Suite du chapitre

Le grand avantage des arcs paramétrés par rapports aux graphes de fonctions est qu'un arc paramétré peut « faire des retours en arrière », ou bien des auto intersections. Outre les deux exemples typiques de la figure 21.4, un exemple classique est la droite verticale. Les fonctions  $y = ax + b$  permettent de décrire toutes les droites, sauf les droites verticales. Dans le cadre des courbes paramétrées, les droites verticales et horizontales sont sur pied d'égalité. Quelques exemples classiques :

**Droite horizontale** Une droite horizontale à la hauteur  $a$  est donnée par la courbe paramétrée  $\gamma(t) = (t, a)$ , avec  $t \in I = \mathbb{R}$ .

**Droite verticale** Une droite verticale à la distance  $b$  de l'origine est donnée par la courbe paramétrée  $\gamma(t) = (b, t)$ , avec  $t \in I = \mathbb{R}$ .

**Graphe d'une fonction** Le graphe d'une fonction  $f: \mathbb{R} \rightarrow \mathbb{R}$  est donné par l'arc  $\gamma(t) = (t, f(t))$ .

**Un cercle** Le cercle de rayon  $R$  est donné par l'arc  $\gamma(t) = (R \cos(t), R \sin(t))$ .

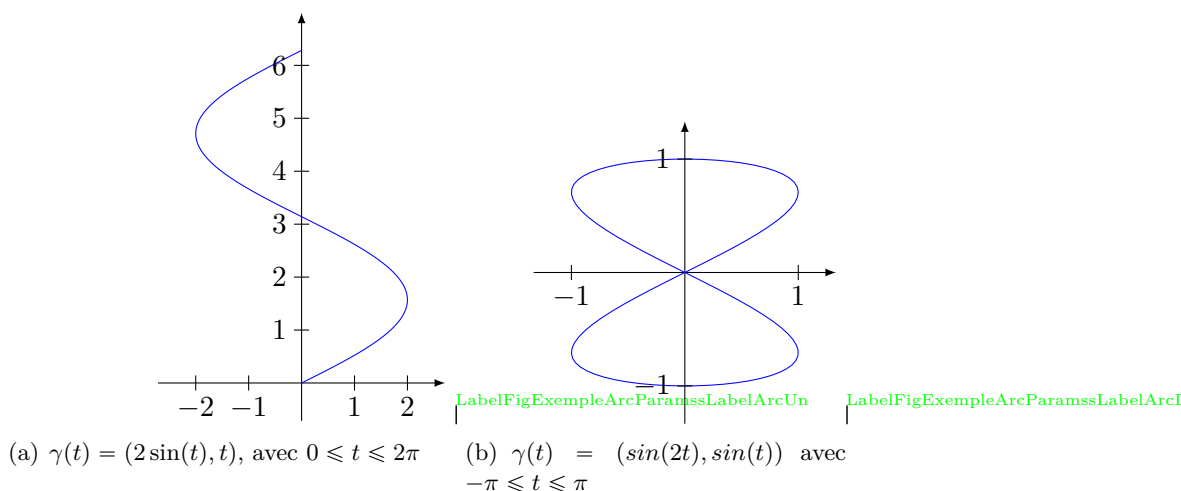


FIGURE 21.4: Des exemples d'arcs paramétrés. Ceux ne sont pas des graphes.

### Remarque 21.20.

Afin d'alléger la notation, nous allons le plus souvent désigner l'arc  $(I, \gamma)$  simplement par la fonction  $\gamma$ . Il est cependant toujours *très* important de savoir sur quel intervalle nous considérons le chemin. Cela dépendra le plus souvent du contexte, et nous indiquerons l'intervalle  $I$  explicitement lorsqu'une ambiguïté est à craindre.

Par exemple, lorsque nous considérons le cercle  $\gamma(t) = (R \cos(t), R \sin(t))$ , le plus souvent l'intervalle de variation de  $t$  sera  $I = [0, 2\pi]$ . Par contre, si nous considérons la droite  $\gamma(t) = (t, 2t)$ , l'intervalle de variation de  $t$  sera naturellement  $I = \mathbb{R}$ .

## 21.5 Autres exemples

### Exemple 21.21.

Soit  $v \in \mathbb{R}^3$  et  $x_0 \in \mathbb{R}^3$ . Le chemin

$$\sigma(t) = x_0 + tv \quad (21.71)$$

est une droite. Sa vitesse est  $\sigma'(t) = v$ . △

### Exemple 21.22.

La courbe

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \in \mathbb{R}^2 \quad (21.72)$$

avec  $t \in [0, 2\pi[$  est le cercle unité parcouru une fois dans le sens trigonométrique.

Notez que si on prend  $t \in [0, 4\pi[$ , nous avons un *autre* chemin ; c'est le même cercle unité, mais parcouru *deux* fois. Même si le « dessin » (le graphe) des deux est le même, le chemin n'est pas le même.

Le chemin

$$\gamma(t) = \begin{pmatrix} \cos(2\pi - t) \\ \sin(2\pi - t) \end{pmatrix} \quad (21.73)$$

est le cercle unité parcouru une fois dans le sens inverse. Encore une fois le « dessin » est le même, mais le chemin n'est pas le même. △

### Exemple 21.23.

Le chemin

$$\sigma(t) = \begin{pmatrix} t \\ t^2 \end{pmatrix} \quad (21.74)$$

est un chemin dont l'image est la parabole d'équation  $y = x^2$ . △

L'importance de la dérivée du chemin réside en le fait qu'elle donne la tangente. En effet le vecteur  $\sigma'(t)$  est tangent au graphe de  $\sigma$  au point  $\sigma(t)$ .

CorKBEMooRvYAcJ

### Corolaire 21.24.

*La tangente à un cercle est perpendiculaire au rayon.*

*Démonstration.* Nous savons que pour un cercle,

$$y'(x) = \frac{-x}{\sqrt{R^2 - x^2}}. \quad (21.75)$$

Un point général du cercle a pour abscisse  $x = R \cos(\theta)$ . En remplaçant nous trouvons le coefficient directeur suivant pour la tangente :

$$y'(R \cos(\theta)) = -\frac{1}{\tan(\theta)}. \quad (21.76)$$

Par conséquent une droite perpendiculaire à la tangente aurait comme coefficient directeur le nombre  $\tan(\theta)$ . Or cela est bien le coefficient directeur du rayon qui joint le point  $(0, 0)$  au point  $(R \cos(\theta), R \sin(\theta))$ . □

### Exemple 21.25.

Pour le cercle,

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad (21.77)$$

la dérivée est donnée par

$$\sigma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \quad (21.78)$$

Le produit scalaire  $\sigma(t) \cdot \sigma'(t)$  est nul. Le vecteur  $\sigma'(t)$  est donc bien tangent (corolaire 21.24).  $\triangle$

**Exemple 21.26.**

Le courbe donnée par le chemin

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix} \quad (21.79)$$

est une hélice. Sa vitesse est

$$\sigma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 1 \end{pmatrix}. \quad (21.80)$$

Notez que pour tout  $t \in \mathbb{R}$ , nous avons  $\|\sigma'(t)\| = \sqrt{2}$ .  $\triangle$

**Remarque 21.27.**

Lorsqu'on parle d'une courbe dans l'espace, l'intervalle sur lequel on considère la variation du paramètre est une donnée fondamentale. Elle fait partie intégrante de la définition de la courbe.

## 21.6 Élément de longueur

### 21.6.1 Élément de longueur : cartésiennes

Étant donné que la longueur d'arc d'une courbe paramétrée  $(I, \gamma)$  est donnée par l'intégrale de  $\|\gamma'(t)\|$ , il est naturel d'appeler le nombre  $\|\gamma'(t)\| dt$ , **l'élément de longueur** de la courbe  $\gamma$  au point  $\gamma(t)$ .

En coordonnées cartésiennes dans le plan, une courbe paramétrée est donnée par

$$\gamma(t) = (x_1(t), x_2(t)), \quad (21.81)$$

et l'élément de longueur est

$$\|x'(t)\| dt = \sqrt{(x_1')^2 + (x_2')^2} dt. \quad (21.82) \quad \text{EqElLongCart}$$

### 21.6.2 Élément de longueur : polaires (1)

En coordonnées polaires, une courbe est donnée par

$$\gamma(t) = (\rho(t), \theta(t)), \quad (21.83)$$

et le passage aux cartésiennes se fait via les formules

$$\begin{cases} x(t) = \rho(t) \cos(\theta(t)) \\ y(t) = \rho(t) \sin(\theta(t)) \end{cases} \quad \begin{matrix} (21.84a) \\ (21.84b) \end{matrix}$$

L'élément de longueur se trouve directement en remplaçant  $x(t)$  et  $y(t)$  dans la formule (21.82). Les dérivées sont données par

$$\begin{aligned} x'(t) &= \rho'(t) \cos \theta(t) - \rho(t) \theta'(t) \sin \theta(t) \\ y'(t) &= \rho'(t) \sin \theta(t) + \rho(t) \theta'(t) \cos \theta(t), \end{aligned} \quad (21.85)$$

et un calcul montre que

$$(x'(t))^2 + (y'(t))^2 = (\rho'(t))^2 + (\rho(t))^2 (\theta'(t))^2. \quad (21.86) \quad \text{EqElLongEnPolaires}$$

Nous reviendrons plus en détail sur le concept de changement de paramétrage (ici, les polaires) à la section 21.7.

### 21.6.3 Élément de longueur : polaires (2)

Parfois on utilise  $\theta$  comme paramètre. L'équation de la courbe est alors donnée en coordonnées polaires sous la forme

$$\rho(\theta) = f(\theta), \quad \text{EqEgenereformepolaire} \quad (21.87)$$

où  $f$  est une fonction réelle et il faut comprendre que nous parlons de la courbe  $(\rho(\theta), \theta)$  en coordonnées polaires. En coordonnées cartésiennes, cette courbe est donnée par |EqPolaireSemiGen

$$\begin{cases} x(t) = \rho(t) \cos(t) \\ y(t) = \rho(t) \sin(t) \end{cases} \quad (21.88a)$$

$$\quad (21.88b)$$

avec  $t$  qui parcourt le plus souvent l'intervalle  $[0, 2\pi]$ . Notez qu'il se peut que le domaine ne soit pas toujours  $[0, 2\pi]$ ; cela peut dépendre des circonstances. Quoiqu'il en soit, la donnée du domaine fait partie de la donnée d'une courbe, et il ne peut donc pas y avoir d'équivoques à ce niveau.

Nous utilisons à nouveau la formule (21.82) en y mettant les valeurs (21.88) :

$$\begin{cases} x'(t) = \rho'(t) \cos(t) - \rho(t) \sin(t) \\ y'(t) = \rho'(t) \sin(t) + \rho(t) \cos(t), \end{cases} \quad (21.89a)$$

$$\quad (21.89b)$$

et

$$(x'(t))^2 + (y'(t))^2 = \rho'(t)^2 + \rho(t)^2. \quad \text{EqElemOngPOldoux} \quad (21.90)$$

#### Remarque 21.28.

N'oubliez pas, en utilisant ces formules, que ce qui rentre dans l'intégrale est la *racine carré* de  $(x')^2 + (y')^2$ .

ExempleLongCercle

#### Exemple 21.29.

Calculons la circonférence du cercle. En coordonnées polaires, le graphe du cercle correspond à l'équation

$$(\rho(t), \theta(t)) = (R, t) \quad (21.91)$$

où  $R$  est constante (le rayon du cercle) et  $t$  va de 0 à  $2\pi$ . En substituant dans l'équation (21.86), l'élément de longueur à intégrer est seulement

$$\sqrt{R^2} = R \quad (21.92)$$

parce que  $\rho'(t) = 0$  et  $\theta'(t) = 1$ . La longueur du cercle est alors directement donnée par

$$l = \int_0^{2\pi} R dt = 2\pi R. \quad (21.93)$$

Nous pouvions aussi faire le calcul en coordonnées cartésiennes. Alors la courbe est donnée par les équations

$$\begin{cases} x(t) = R \cos(t) \\ y(t) = R \sin(t) \end{cases} \quad (21.94)$$

et  $t \in [0, 2\pi]$ . La circonférence du cercle est alors

$$l = \int_0^{2\pi} \sqrt{R^2 \sin^2(t) + R^2 \cos^2(t)} dt = \int_0^{2\pi} R dt = 2\pi R. \quad (21.95)$$

△

#### Remarque 21.30.

Il faut bien comprendre que quand on parle de courbes paramétrées en coordonnées cartésiennes on pense à une courbe dont le paramètre est, par exemple,  $t$  et les équations de la courbe sont  $(x(t), y(t))$ . Cela ne veut pas dire que  $x$  ou  $y$  soit le paramètre. Le cas où  $x$  ou  $y$  est le paramètre est un cas particulier qui est possible seulement pour certaines courbes et notamment pour les graphes. Le cercle de rayon 1 n'est pas un graphe, donc si on veut utiliser  $x$  ou  $y$  comme paramètre il faut d'abord découper la courbe en deux morceaux, par exemple, la moitié inférieure ( $y < 0$ ) et la moitié supérieure ( $y > 0$ ).

ExCycloLong

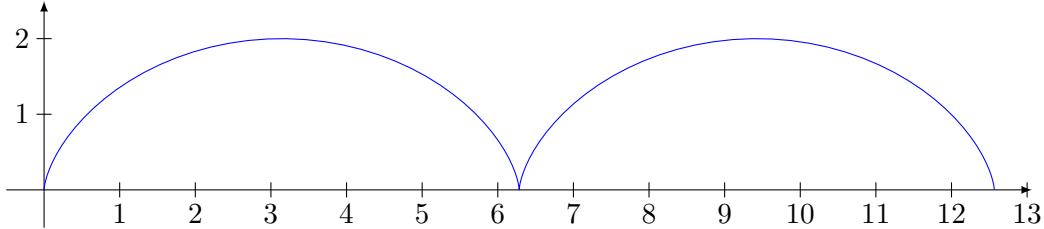
**Exemple 21.31.**

Une **cycloïde** est une courbe paramétrée par

$$\begin{cases} x(t) = a(t - \sin(t)) \\ y(t) = a(1 - \cos(t)) \end{cases} \quad (21.96a)$$

$$(21.96b)$$

avec  $a > 0$  et  $t \in \mathbb{R}$ . Comme montré sur la figure 21.5, la cycloïde donne lieu à un graphe périodique. Il est possible de montrer (le faire) que le premier arc correspond à  $t \in [0, 2\pi]$ . Nous voulons donc calculer la longueur de l'arc sur cet intervalle.

FIGURE 21.5: La cycloïde de paramètre  $a = 1$  entre 0 et  $4\pi$ .

LabelFigCycloïdeA

Nous avons  $x'(t) = a(1 - \cos(t))$  et  $y'(t) = a \sin(t)$ , de telle façon que

$$\sqrt{(x')^2 + (y')^2} = a\sqrt{2 - 2\cos(t)} = a\sqrt{4\sin^2\left(\frac{t}{2}\right)} = 2a\left|\sin\frac{t}{2}\right|. \quad (21.97)$$

Eq\_0508dlcycloïde

La longueur est donc donnée par

$$\int_0^{2\pi} 2a\left|\sin\frac{t}{2}\right|dt = 4a \int_0^{\pi} \sin(t)dt = 8a. \quad (21.98)$$

△

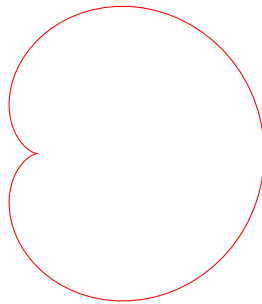
**Exemple 21.32.**

La **cardioïde** est la courbe donnée par

$$\rho(\theta) = a(1 + \cos(\theta)).$$

EqCardioïde (21.99)

avec  $\theta \in [-\pi, \pi]$ . Le nom de cette courbe provient de son graphe illustré à la figure 21.6.

FIGURE 21.6: Une cardioïde,  $\rho = 1 + \cos(\theta)$ .

LabelFigCardioid

L'équation (21.99) est donnée sous la forme (21.87), c'est-à-dire que  $\theta(t) = t$  et  $\theta'(t) = 1$ , et par conséquent l'élément de longueur est donné par

$$\begin{aligned} (\rho')^2 + (\rho)^2 &= (-a \sin(\theta))^2 + a^2(1 + \cos(\theta))^2 \\ &= a^2 \sin^2(\theta) + a^2(1 + 2\cos(\theta) + \cos^2(\theta)) \\ &= a^2(1 + 1 + 2\cos(\theta)) \\ &= 2a^2(1 + \cos(\theta)) \\ &= 4a^2 \cos^2 \frac{\theta}{2}. \end{aligned} \quad (21.100)$$



La longueur d'arc est donc donnée par

$$l = \int_{-\pi}^{\pi} 2a \cos \frac{\theta}{2} d\theta = 2a \int_{-\pi/2}^{\pi/2} \cos(t) 2dt = 8a. \quad (21.101)$$

△

#### 21.6.4 Approximation de la longueur par des cordes

**Définition 21.33** (Point régulier, birégulier[669]).

Soit un arc paramétré  $(I, \gamma)$ . Un point  $t \in I$  est dit **régulier** si  $\gamma'(t) \neq 0$ , et il est dit **critique** si  $\gamma'(t) = 0$ . Le point  $t \in I$  est dit **birégulier** si les vecteurs  $\gamma'(t)$  et  $\gamma''(t)$  sont linéairement indépendants et non nuls.

Par extension, nous dirons également que le point  $\gamma(t)$  lui-même est régulier, critique ou birégulier. Un arc est dit régulier lorsque tous ses points sont réguliers.

Note : dans le lemme 21.71 et ses dépendances, nous utilisons effectivement que l'arc  $\gamma$  est de classe  $C^2$ .

Nous savons que la longueur d'une courbe est donnée par le supremum sur toutes les subdivisions de la longueur des cordes correspondantes. De plus l'inégalité triangulaire nous enseigne que plus la subdivision est fine, plus la longueur sera grande. Il est donc naturel de penser que sur un petit intervalle, la longueur de la courbe ne doit pas être très différente de la longueur de la corde correspondante.

La proposition suivante est un énoncé précis et quantitatif de ce fait.

**Proposition 21.34.**

Soit  $(I, \gamma)$  un arc de classe  $C^1$  et  $t_0 \in I$  un point régulier (c'est-à-dire  $\gamma'(t_0) \neq 0$ ). Alors pour tout  $\varepsilon > 0$ , il y a un  $\delta > 0$  tel que on trouve  $t, t' \in I \cap (t_0, \delta)$  tels que

$$\left| \int_t^{t'} \|\gamma'(u)\| du - \|\gamma(t) - \gamma(t')\| \right| \leq 2\varepsilon |t - t'|. \quad (21.102)$$

Intuitivement, cette proposition signifie qu'au voisinage de  $t_0$ , la longueur d'arc est équivalente à celle de la corde.

*Démonstration.* Par la continuité de  $\gamma'$  (parce que  $\gamma$  est  $C^1$ ), pour tout  $\varepsilon$ , il existe un  $\delta$  tel que

$$|t - t_0| < \delta \Rightarrow \|\gamma'(t) - \gamma'(t_0)\| \leq \varepsilon. \quad (21.103)$$

Nous considérons la fonction

$$u \mapsto \gamma(u) - \gamma(t_0) - (u - t_0)\gamma'(t_0), \quad (21.104)$$

dont la dérivée (par rapport à  $u$ ) est

$$\gamma'(u) - \gamma'(t_0). \quad (21.105)$$

Nous y appliquons la formule des accroissements finis entre  $t$  et  $t'$  choisis dans  $I \cap ]t_0 - \delta, t_0 + \delta[$ . Il existe un  $u$  entre  $t$  et  $t'$  tel que

$$\begin{aligned} \|\gamma(t) - \gamma(t_0) - (t - t_0)\gamma'(t_0) - \gamma(t') + \gamma(t_0) + (t' - t_0)\gamma'(t_0)\| \\ = |t - t'| \|\gamma'(u) - \gamma'(t_0)\| \\ \leq \varepsilon |t - t'|. \end{aligned} \quad (21.106)$$

En simplifiant ce qui peut être simplifié dans le membre de gauche, nous trouvons

$$\|\gamma(t) - \gamma(t') - (t - t')\gamma'(t_0)\| \leq \varepsilon |t - t'|. \quad (21.107)$$

Le membre de gauche peut être minoré en utilisant la proposition 7.213 :

$$\left| \|\gamma(t) - \gamma(t')\| - \|(t - t')\gamma'(t_0)\| \right| \leq \varepsilon |t - t'|. \quad (21.108)$$

D'autre part, les inégalités (7.195) montrent que

$$-\|\gamma'(u) - \gamma'(t_0)\| \leq \|\gamma'(u)\| - \|\gamma'(t_0)\| \leq \|\gamma'(u) - \gamma'(t_0)\|. \quad (21.109)$$

Si de plus  $u$  est compris entre  $t$  et  $t'$ , ces inégalités sont encore coincées entre  $-\varepsilon$  et  $\varepsilon$ . En intégrant (21.109) par rapport à  $u$  entre  $t$  et  $t'$ , nous obtenons

$$\left| \int_t^{t'} \|\gamma'(u)\| - (t - t')\|\gamma'(t_0)\| \right| \leq \varepsilon |t - t'|. \quad (21.110)$$

Afin d'alléger les notations pour la ligne suivante, nous notons  $A$  le nombre positif  $\int_t^{t'} \|\gamma'(u)\| du$ . Nous avons

$$\begin{aligned} \left| A - \|\gamma(t) - \gamma(t')\| \right| &= \left| A - |t - t'| \|\gamma'(t_0)\| + |t - t'| \|\gamma'(t_0)\| - \|\gamma(t) - \gamma(t')\| \right| \\ &\leq \left| A - |t - t'| \|\gamma'(t_0)\| \right| + \left| |t - t'| \|\gamma'(t_0)\| - \|\gamma(t) - \gamma(t')\| \right|. \end{aligned} \quad (21.111)$$

L'équation (21.108) montre que le second terme est plus petit ou égal à  $\varepsilon |t - t'|$ . En ce qui concerne le premier terme, étant donné que  $A$  est positif,

$$\left| A - |t - t'| \|\gamma'(t_0)\| \right| \leq \left| A - (t - t') \|\gamma'(t_0)\| \right| \leq \varepsilon |t - t'|. \quad (21.112)$$

Au final, l'inéquation (21.111) donne

$$\left| A - \|\gamma(t) - \gamma(t')\| \right| \leq 2\varepsilon |t - t'|, \quad (21.113)$$

ce qu'il fallait démontrer.  $\square$

## 21.7 Arc géométrique

### Définition 21.35.

Soient  $(I, \gamma)$  et  $(J, g)$  deux arcs paramétrés de classe  $\mathcal{C}^k$ . On dit qu'ils sont **équivalents** si il existe une bijection  $\theta: I \rightarrow J$  de classe  $\mathcal{C}^k$ , d'inverse de classe  $\mathcal{C}^k$  telle que  $g = \gamma \circ \theta$ . Nous notons  $\gamma \sim g$  lorsque  $\gamma$  et  $g$  sont équivalents (les ensembles  $I$  et  $J$  sont sous-entendus).

Le passage d'un paramétrage  $(I, \gamma)$  à une autre  $(J, g)$  se fait selon le diagramme suivant :

$$\begin{array}{ccc} I & \xrightarrow{\gamma} & \mathbb{R}^n \\ \uparrow \theta & \nearrow g & \\ J & & \end{array} \quad (21.114)$$

### Proposition 21.36.

La relation donnée dans la définition 21.35 est une relation d'équivalence.

**Démonstration.** Les trois points d'une relation d'équivalence se vérifient en utilisant le fait que  $\theta$  est inversible, et que l'inverse  $\theta^{-1}$  jouit des mêmes propriétés de continuité ( $\mathcal{C}^k$ ) que  $\theta$ .

**Réflexivité** Nous avons  $\gamma \sim \gamma$  avec  $\theta = \text{Id}$ .

**Symétrie** Si  $\gamma \sim g$ , alors nous avons une application  $\theta$  telle que  $g = \gamma \circ \theta$ , et donc  $\gamma = g \circ \theta^{-1}$ , ce qui montre que  $g \sim \gamma$ .

**Transitivité** Si  $\gamma \sim g$  et  $g \sim h$  avec  $g = \gamma \circ \theta$  et  $h = g \circ \omega$ , alors  $h = \gamma \circ (\theta \circ \omega)$ , ce qui montre que  $\gamma \sim h$ .

□

Si les arcs  $(I, \gamma)$  et  $(J, g)$  sont équivalents, les images dans  $\mathbb{R}^n$  sont identiques, et décrivent donc « le même dessin ». Nous allons préciser cette notion plus loin.

**Définition 21.37.**

Pour cette raison les classes d'équivalences sont appelées des **arcs géométriques** (de classe  $\mathcal{C}^k$ ).

Si  $\Gamma$  est un arc géométrique, ses représentants sont dits des **paramétrages admissibles** ou, plus simplement *paramétrage*. On dit que l'application  $\theta: J \rightarrow I$  est un **changement de variable**. Nous disons que un arc géométrique est *compact* quand ses représentants sont compacts.

LemChamVarsStriMomnot

**Lemme 21.38.**

Dans le cas d'un arc  $\mathcal{C}^1$ , les changements de variables sont strictement monotones (croissants ou décroissants).

*Démonstration.* Nous considérons  $(I, \gamma)$  et  $(J, g)$ , deux paramétrages différents du même arc géométrique, et  $\theta \in \mathcal{C}^1(J, I)$  le changement de variable. Nous allons noter  $t$  la variable sur  $I$  et  $s$  la variable sur  $J$ . Par définition,  $\theta(\theta^{-1}(t)) = t$ , et par conséquent,

$$\theta'(\theta^{-1}(t))(\theta^{-1})'(t) = 1. \quad (21.115)$$

En particulier  $\theta'(\theta^{-1}(t))$  ne s'annule pas pour aucune valeur de  $t$ . Mais  $\theta^{-1}(t)$  peut prendre n'importe quelle valeur dans  $J$ , donc nous avons  $\theta'(s) \neq 0$  pour tout  $s \in J$ . Cela signifie bien que  $\theta$  est strictement monotone. En effet,  $\theta'$  étant continue, elle ne peut pas changer de signe sans passer par zéro (théorème 9.95 des valeurs intermédiaires). □

ThoLongArcGeom

**Théorème 21.39.**

La longueur d'un arc est indépendante de son paramétrage, c'est-à-dire que les représentants d'un arc géométrique compact de classe  $\mathcal{C}^1$  ont même longueur.

*Démonstration.* Nous utilisons les mêmes notations que celles du lemme 21.38. Nous savons déjà que le changement de variable  $\theta: J \rightarrow I$  est strictement monotone. Supposons que  $\theta$  soit croissante. En effectuant un changement de variable dans l'intégrale qui donne la longueur<sup>6</sup> nous avons

$$\begin{aligned} l(\gamma) &= \int_I \|\gamma'(t)\| dt \\ &= \int_J \|\gamma'(\theta(s))\| \theta'(s) ds \\ &= \int_J \|\gamma'(\theta(s))\theta'(s)\| ds \\ &= \int_J \left\| \frac{d}{ds} (\gamma \circ \theta)(s) \right\| ds \\ &= \int_J \|g'(s)\| ds \\ &= l(J, g). \end{aligned} \quad (21.116)$$

□

**Définition 21.40.**

Nous nommons **longueur** d'un arc géométrique la longueur commune de tous ses représentants. On dit que l'arc géométrique est **rectifiable** si sa longueur est  $< \infty$ .

---

6. Théorème 21.9.

### 21.7.1 Abscisse curviligne et paramétrage normal

SubSecAbsCurv

#### Définition 21.41.

Soit  $(I, \gamma)$  un arc paramétré continu rectifiable. Nous appelons **abscisse curviligne** de  $\gamma$  toute application  $\phi: I \rightarrow \mathbb{R}$  telle que pour tout  $t, t' \in I$  avec  $t < t'$ , nous ayons

$$l([t, t'], \gamma) = |\phi(t') - \phi(t)|. \quad (21.117)$$

Si il existe un  $t_0 \in I$  tel que  $\phi(t_0) = 0$ , alors nous disons que  $t_0$  est l'**origine** de l'abscisse  $\phi$ .

DEFooJJQFooEITCvG

#### Définition 21.42.

Un arc paramétré  $(I, \gamma_N)$  continu rectifiable est dit **normal** si l'identité est une abscisse curviligne.

LEMooLADUooBIHjuT

#### Lemme 21.43.

Si  $\gamma$  est de classe  $C^1$  et est un paramétrage normal, alors

(1) pour tout choix de  $t$  et  $t'$  dans  $I$  avec  $t < t'$ , nous avons

$$l([t, t'], \gamma_N) = t' - t. \quad (21.118)$$

(2)  $\|\gamma'(t)\| = 1$  pour tout  $t$ .

*Démonstration.* Pour tout  $x_1, x_2$  dans le domaine nous avons

$$l([x_1, x_2], \gamma) = \int_{x_1}^{x_2} \|\gamma'(t)\| dt = x_2 - x_1. \quad (21.119)$$

Cela implique  $\|\gamma'(t)\| = 1$  pour tout  $t$ . En effet, pour fixer les idées, supposons que  $\|\gamma'(t)\| > 1$  en un point, par continuité, cela reste strictement supérieur à 1 sur un intervalle<sup>7</sup>. L'intégrale sur cet intervalle ne peut alors pas être la taille de l'intervalle.  $\square$

ExCerlceRadNorm

#### Exemple 21.44.

Le cercle unitaire est donné par l'arc

$$\gamma(t) = (\cos(t), \sin(t)) \quad (21.120)$$

et  $t \in [0, 2\pi]$ . Pour tout choix de  $t$  et  $t'$  dans  $[0, 2\pi]$ , nous avons

$$l([t, t'], \gamma) = \int_t^{t'} \sqrt{\sin^2(u) + \cos^2(u)} du = t' - t. \quad (21.121)$$

Les angles exprimés en radians forment donc un paramétrage normal du cercle de rayon 1.  $\triangle$

#### Lemme 21.45.

Pour un arc paramétré compact, la longueur d'arc est une abscisse curviligne.

*Démonstration.* Par définition de la longueur d'arc  $\varphi$ , nous avons

$$\varphi(t') - \varphi(t) = l([a, t'], \gamma) - l([a, t], \gamma) = \diamond. \quad (21.122)$$

Supposons pour fixer les idées que  $t' > t$ . En utilisant la proposition 21.6, nous avons

$$l([a, t'], \gamma) = l([a, t], \gamma) + l([t, t'], \gamma), \quad (21.123)$$

et donc après simplification de deux termes,

$$\diamond = l([t, t'], \gamma), \quad (21.124)$$

ce qui est précisément la propriété demandée pour être une abscisse curviligne.  $\square$

7. Corolaire 9.91.

PropExisteChmNorm

**Proposition 21.46.**

Pour tout arc paramétré  $C^1$  sans points critiques, il existe un changement de coordonnées qui rend l'arc normal.

*Démonstration.* Soit  $(I, \gamma)$  un arc de classe  $C^1$ . Nous devons montrer qu'il existe un intervalle  $J$  et une application  $\theta: J \rightarrow I$  de classe  $C^1$  et d'inverse  $C^1$  tel que l'arc  $(J, \gamma_N)$  soit  $C^1$  où  $\gamma_N = \gamma \circ \theta$ .

Si  $I = [a, b]$ , nous considérons la fonction

$$\begin{aligned} \phi: I &\rightarrow \mathbb{R}^+ \\ t &\mapsto \int_a^t \|\gamma'(u)\| du. \end{aligned} \quad \text{EqDevVarPhi} \quad (21.125)$$

Étant définie par l'intégrale d'une fonction  $C^0$ , la fonction  $\phi$  est  $C^1$ , et nous avons  $\phi'(t) = \|\gamma'(t)\| > 0$  pour tout  $t \in I$ . Vue comme application  $\phi: [a, b] \rightarrow [0, l(\gamma)]$ , l'application  $\phi$  est bijective et d'inverse  $C^1$ . Voyons cela point par point.

- (1) La fonction  $\phi$  est injective parce qu'elle est strictement croissante.
- (2) Elle est surjective parce que  $\phi(a) = 0$  et  $\phi(b) = l(\gamma)$ .
- (3) La continuité de l'inverse est plus délicate. Soit  $s \in [0, l(\gamma)]$  et  $\varepsilon > 0$ . Pour prouver la continuité de  $\phi^{-1}$  en  $s$ , nous devons trouver un  $\delta$  tel que

$$|s - s'| < \delta \Rightarrow |\phi^{-1}(s) - \phi^{-1}(s')| < \varepsilon. \quad (21.126)$$

Étant donné que  $s$  et  $s'$  sont dans l'image de  $\phi$ , nous considérons les uniques  $t$  et  $t'$  tels que  $s = \phi(t)$  et  $s' = \phi(t')$ . La quantité  $\phi(t) - \phi(t')$  devient

$$\int_a^t \|\gamma'(u)\| du - \int_a^{t'} \|\gamma'(u)\| du = \int_t^{t'} \|\gamma'(u)\| du. \quad \text{EqCondypemuCont} \quad (21.127)$$

D'autre part,  $\phi^{-1}(s) = t$  et  $\phi^{-1}(s') = t'$ , donc la condition (21.127) devient

$$\left| \int_t^{t'} \|\gamma'(u)\| du \right| \leq \delta \Rightarrow |t - t'| < \varepsilon. \quad (21.128)$$

Cela revient à la continuité des fonctions définies par une intégrale.

- (4) La dérivée de son inverse est donnée par <sup>8</sup>

$$(\phi^{-1})'(s) = \frac{1}{\phi'(\phi^{-1}(s))}. \quad (21.129)$$

Nous avons vu que  $\phi^{-1}$  et  $\phi'$  étaient continues. La fonction  $(\phi^{-1})'$  étant exprimée en termes de ces deux fonctions elle est également continue.

Nous considérons l'arc paramétré  $(J, \gamma_N)$  avec  $J = [0, l(\gamma)]$  et

$$\gamma_N(s) = (\gamma \circ \phi^{-1})(s). \quad (21.130)$$

---

8. Pour obtenir cette formule, dérivez les deux membres de l'équation  $\phi(\phi^{-1}(s)) = s$ .

Nous montrons maintenant que ce nouveau paramétrage est normal. Soient  $0 \leq s \leq s' \leq l(\gamma)$ ,

$$\begin{aligned}
 l([s, s'], g) &= \int_s^{s'} \|\gamma'_N(u)\| du \\
 &= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|(\gamma'_N \circ \phi)(t)\| \phi'(t) dt \\
 &= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|(\gamma_N \circ \phi)'(t)\| dt \\
 &= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|\gamma'(t)\| dt \\
 &= \int_0^{\phi^{-1}(s')} \|\gamma'(t)\| dt - \int_0^{\phi^{-1}(s)} \|\gamma'(t)\| dt \\
 &= \phi(\phi^{-1}(s')) - \phi(\phi^{-1}(s)) \\
 &= s' - s,
 \end{aligned} \tag{21.131}$$

ce qui prouve que le paramétrage  $(J, \gamma_N)$  est normal.  $\square$

Nous retenons que le paramétrage normal de  $\gamma$  est donné par  $(J, \gamma_N)$  avec  $J = [0, l(\gamma)]$  et

$$\gamma_N(s) = (\gamma \circ \phi^{-1})(s) \tag{21.132}$$

où

$$\begin{aligned}
 \phi: I &\rightarrow \mathbb{R}^+ \\
 t &\mapsto \int_a^t \|\gamma'(u)\| du.
 \end{aligned} \tag{21.133}$$

Notons aussi que  $\phi$  est une fonction croissante, étant l'intégrale d'une fonction positive.

### Exemple 21.47.

Trouvons les coordonnées normales pour la cycloïde donnée par

$$\begin{cases} x(t) = a(t - \sin(t)), \\ y(t) = a(1 - \cos(t)) \end{cases} \tag{21.134a}$$

$$\tag{21.134b}$$

et  $t \in ]0, 2\pi[$ . Relire l'exemple 21.31.

D'abord nous trouvons  $\phi$  avec la formule (21.133) avec  $a = 0$ . En utilisant le bout de calcul (21.97), nous avons

$$\phi(t) = 2a \int_0^t \sin \frac{u}{2} du = 4a \left(1 - \cos \frac{t}{2}\right). \tag{21.135}$$

Pour trouver  $\phi^{-1}(s)$ , nous résolvons l'équation

$$s = \phi(\phi^{-1}(s)) \tag{21.136}$$

par rapport à  $\phi^{-1}(s)$ . Dans un premier temps, nous trouvons

$$1 - \frac{s}{4a} = \cos \frac{\phi^{-1}(s)}{2}, \tag{21.137}$$

donc  $\frac{\phi^{-1}(s)}{2} = \arccos(\frac{4a-s}{4a})$ , et finalement

$$\phi^{-1}(s) = 2 \arccos \left( \frac{4a-s}{4a} \right). \tag{21.138}$$

Il nous reste à injecter cela dans les expressions de  $x(t)$  et  $y(t)$  pour trouver  $(\gamma_N)_x(s)$  et  $(\gamma_N)_y(s)$ . D'abord,

$$(\gamma_N)_x(s) = a[\phi^{-1}(s) - \sin(\phi^{-1}(s))]. \tag{21.139}$$

Nous utilisons maintenant la formule trigonométrique  $\sin(x) = 2 \sin \frac{x}{2} \cos \frac{x}{2}$  afin de simplifier les expressions :

$$\begin{aligned}
 (\gamma_N)_x &= a \left[ 2 \arccos \left( \frac{4a-s}{4a} \right) - 2 \sin \left( \arccos \left( \frac{4a-s}{4a} \right) \right) \cos \left( \arccos \left( \frac{4a-s}{4a} \right) \right) \right] \\
 &= a \left[ 2 \arccos \left( \frac{4a-s}{4a} \right) - \frac{4a-s}{2a} \sqrt{1 - \left( \frac{4a-s}{4a} \right)^2} \right] \\
 &= 2a \arccos \left( \frac{4a-s}{4a} \right) - \sqrt{8as - s^2} \frac{4a-s}{8a}
 \end{aligned} \tag{21.140}$$

où nous avons utilisé la formule  $\sin(\arccos(x)) = \sqrt{1-x^2}$ . Ensuite, pour obtenir  $(\gamma_N)_y$  nous devons calculer

$$(\gamma_N)_y(s) = a[1 - \cos(\phi^{-1}(s))]. \tag{21.141}$$

Encore une fois, il est intéressant d'exprimer le cosinus en termes des angles divisés par deux :  $\cos(x) = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}$ .

$$\begin{aligned}
 (\gamma_N)_y &= a \left[ 1 - \cos^2 \frac{\phi^{-1}(s)}{2} + \sin^2 \frac{\phi^{-1}(s)}{2} \right] \\
 &= a \left[ 2 - 2 \cos^2 \frac{\phi^{-1}(s)}{2} \right] \\
 &= 2a \left[ 1 - \left( \frac{4a-s}{4a} \right)^2 \right].
 \end{aligned} \tag{21.142}$$

Dans ce paramétrage,  $s \in ]0, 8a[$ .

△

### Exemple 21.48.

La cardioïde  $\rho(\theta) = a(1 + \cos(\theta))$  avec  $\theta$  entre  $-\pi$  et  $\pi$ . Avant d'utiliser la formule (21.133), nous devons trouver l'élément de longueur de la cardioïde. Étant donné la façon dont l'équation de la cardioïde nous est donnée, l'élément de longueur est donné par <sup>9</sup> (21.90) :

$$\begin{aligned}
 \|\gamma'(u)\|^2 &= a^2 \sin^2(u) + a^2(1 + \cos(u))^2 \\
 &= 2a^2(1 + \cos(u)),
 \end{aligned} \tag{21.143}$$

et par conséquent <sup>10</sup>

$$\begin{aligned}
 \phi(t) &= \int_0^t \sqrt{2a^2(1 + \cos(u))} du \\
 &= \int_0^t \sqrt{2a^2 \left( 1 + \cos^2 \frac{u}{2} - \sin^2 \frac{u}{2} \right)} du \\
 &= 2a \int_0^t \cos \frac{u}{2} du \\
 &= 4a \sin \frac{t}{2}.
 \end{aligned} \tag{21.144}$$

Pour trouver l'inverse, nous résolvons  $\phi(\phi^{-1}(s)) = s$  par rapport à  $\phi^{-1}(s)$  :

$$\begin{aligned}
 4a \sin \left( \frac{\phi^{-1}(s)}{2} \right) &= s, \\
 \phi^{-1}(s) &= 2 \arcsin \left( \frac{s}{4a} \right).
 \end{aligned} \tag{21.145}$$

9. Nous vous déconseillons d'étudier cette formule par cœur. Sachez cependant la retrouver assez vite.

10. L'utilisation stricte de la formule (21.133) demanderait d'intégrer à partir de  $-\pi$ . Pour plus de simplicité, nous intégrons à partir de zéro, et nous verrons plus tard comment adapter l'intervalle du nouveau paramètre.

Avant d'écrire trop brutalement  $\gamma_N(s) = (\gamma \circ \phi^{-1})(s)$ , il faut comprendre comment est  $\gamma$ . Nous avons reçu la courbe sous forme polaire, c'est-à-dire

$$\gamma(t) = (\gamma_r(t), \gamma_\theta(t)) = (a(1 + \cos(t)), t). \quad (21.146)$$

C'est comme cela qu'il faut comprendre la donnée  $\rho(\theta) = a(1 + \cos(\theta))$ . Maintenant la formule  $\gamma_N(s) = (\gamma \circ \phi^{-1})(s)$  devient

$$\begin{cases} (\gamma_N)_r(s) = \gamma_r(\phi^{-1}(s)) \\ (\gamma_N)_\theta(s) = \gamma_\theta(\phi^{-1}(s)). \end{cases} \quad \begin{matrix} (21.147a) \\ (21.147b) \end{matrix}$$

Étant donné que  $\gamma_\theta(t) = t$ , la seconde est facile :

$$(\gamma_N)_\theta(s) = 2 \arcsin\left(\frac{s}{4a}\right). \quad (21.148)$$

Pour la première,

$$(\gamma_N)_r(s) = a[1 + \cos(2 \arcsin \frac{s}{4a})] = \frac{16a^2 - s^2}{8a}. \quad (21.149)$$

Nous écrivons donc le nouveau paramétrage en coordonnées polaires sous la forme

$$\left(\frac{16a^2 - s^2}{8a}, 2 \arcsin \frac{s}{4a}\right). \quad (21.150)$$

La question qui arrive maintenant est de savoir quel intervalle parcourt la nouvelle variable  $s$ . D'après le résultat de l'exemple 21.99, la longueur de la cardioïde est de  $8a$  et nous avons donc  $s \in [0, 8a]$ . Cependant, la condition d'existence de arcsin nous interdit d'avoir  $s$  plus grand que  $4a$  en valeur absolue. Où est le problème ?

Le problème est que nous avons changé l'origine de notre paramètre en donnant  $\phi(t)$  comme une intégrale à partir de 0 au lieu de  $-\pi$ . Cela se voit en regardant de quel point nous partons : en  $s = 0$  nous sommes sur le point  $(2a, 0)$  tandis qu'avec le paramètre original, c'est-à-dire  $\theta \in [-\pi, \pi]$ , nous avons pour  $\theta = -\pi$  le point  $(0, -\pi)$ .

Il se passe donc que si nous commençons à parcourir la cardioïde avec  $s = 0$ , nous partons du milieu, et nous ne parcourons donc pas tout. Étant donné que le « premier » point de la cardioïde est le point  $(0, -\pi)$ , le paramètre  $s$  commence en  $s = -4a$ , et nous avons comme intervalle :

$$s \in [-4a, 4a], \quad (21.151)$$

ce qui est en accord avec la conditions d'existence.  $\triangle$

Quel enseignement tirer de cet exemple ? Lorsqu'on calcule  $\phi(t)$  pour trouver les coordonnées normales, il y a deux solutions.

- (1) Utiliser strictement la formule  $\phi(t) = \int_a^t \|\gamma'(u)\| du$ , en prenant bien comme borne de départ le point de départ de le paramétrage de  $\gamma$ . À ce moment la coordonnée normale construite aura  $[0, l(\gamma)]$  comme intervalle de variation.
- (2) Faire commencer l'intervalle d'intégration en zéro (ou ailleurs). Un bon choix peut simplifier quelques calculs, mais alors il faudra bien choisir la valeur de départ de la nouvelle coordonnées pour que le « premier » point de la courbe soit correct. Dans ce cas, la longueur de l'intervalle sera quand même  $l(\gamma)$ . Il n'y a donc pas de problèmes pour trouver la valeur du bout de l'intervalle de variation du paramètre normal.

Dans tous les cas, il faut bien préciser l'intervalle de variation du paramètre lorsqu'on donne une courbe paramétrée.



### 21.7.2 Tangente à une courbe paramétrée

#### Définition 21.49.

Soit  $(I, \gamma)$  un arc paramétré de classe  $\mathcal{C}^k$  avec  $k \geq 1$ . Nous disons que la courbe admet une **tangente** en  $\gamma(t_0) \in \mathbb{R}^n$  lorsque les deux conditions suivantes sont remplies

- (1)  $\gamma(t) \neq \gamma(t_0)$  pour tout  $t$  dans un voisinage de  $t_0$  ;
- (2) la direction de la droite qui passe par  $\gamma(t)$  et  $\gamma(t_0)$  admet une limite lorsque  $t \rightarrow t_0$ .

Dans ce cas, la tangente sera la droite passant par le point  $\gamma(t_0)$  et dont la direction est donnée par la limite.

Dans cette définition, par **direction** d'une droite, nous entendons le vecteur de norme 1 parallèle à celle-ci sans tenir compte du signe. La tangente sera donc la droite passant par  $\gamma(t_0)$  et parallèle au vecteur

$$\lim_{t \rightarrow t_0} \frac{\gamma(t) - \gamma(t_0)}{\|\gamma(t) - \gamma(t_0)\|}. \quad (21.152)$$

Évidemment si nous avons écrit  $\gamma(t_0) - \gamma(t)$ , ça n'aurait pas changé la droite. Par abus de langage, nous parlerons souvent de « la direction  $u$  » même lorsque  $u$  n'est pas de norme 1.

Formellement, une direction est une classe d'équivalence de vecteurs pour la relation  $u \sim v$  si il existe  $\lambda \neq 0$  tel que  $u = \lambda v$ , mais nous n'aurons pas besoin de cette précision ici.

Sans surprises, la tangente est à peu près toujours donnée par la dérivée lorsqu'elle existe. Plus précisément nous avons le

#### Théorème 21.50.

Soit  $(I, \gamma)$ , un arc paramétré de classe  $\mathcal{C}^k$  ( $k \geq 1$ ) et  $t_0 \in I$  tel que

$$\gamma'(t_0) = \gamma''(t_0) = \dots = \gamma^{(q-1)}(t_0) = 0 \quad (21.153)$$

et

$$\gamma^{(q)}(t_0) \neq 0 \quad (21.154)$$

pour un entier  $1 \leq q \leq k$ . Alors  $\gamma$  admet une tangente en  $\gamma(t_0)$  de direction  $\gamma^{(q)}(t_0)$ .

*Démonstration.* Le développement de  $\gamma(t_0)$  en série de Taylor autour de  $t$  jusqu'à l'ordre  $q$  est

$$\begin{aligned} \gamma(t) &= \gamma(t_0) + \gamma'(t_0)|t - t_0| + \frac{\gamma''(t_0)}{2}|t - t_0|^2 + \dots + \frac{\gamma^{(q)}(t_0)}{q!}|t - t_0|^q \\ &\quad + \varepsilon(t)|t - t_0|^q \end{aligned} \quad \text{EqDexTaylor} \quad (21.155)$$

où  $\varepsilon$  est une application  $\varepsilon: \mathbb{R} \rightarrow \mathbb{R}^n$  telle que  $\lim_{t \rightarrow t_0} \varepsilon(t) = 0$ . En utilisant les hypothèses, nous éliminons la majorité des termes dans le développement (21.155) :

$$\gamma(t) - \gamma(t_0) = \frac{1}{q!} \gamma^{(q)}(t_0) |t - t_0|^q + \varepsilon(t) |t - t_0|^q. \quad (21.156)$$

La direction de la droite qui joint  $\gamma(t)$  à  $\gamma(t_0)$  est donc donnée par

$$\frac{\gamma(t) - \gamma(t_0)}{\|\gamma(t) - \gamma(t_0)\|} = \frac{\frac{1}{q!} \gamma^{(q)}(t_0) |t - t_0|^q + \varepsilon(t) |t - t_0|^q}{\|\frac{1}{q!} \gamma^{(q)}(t_0) |t - t_0|^q + \varepsilon(t) |t - t_0|^q\|} \quad (21.157)$$

et la limite lorsque  $t \rightarrow t_0$  donne  $\gamma^{(q)}(t_0)$  comme direction de la tangente. □

Lorsque le théorème s'applique, le vecteur

$$\tau = \frac{\gamma^{(q)}(t_0)}{\|\gamma^{(q)}(t_0)\|} \quad (21.158)$$

est appelé le **vecteur unitaire tangent** en  $\gamma(t_0)$  à l'arc paramétré  $\gamma$ .

CorTgSoCum

**Corolaire 21.51.**

Si  $(I, \gamma)$  est un arc paramétré de classe  $\mathcal{C}^1$  régulier (c'est-à-dire  $\gamma'(t) \neq 0$  pour tout  $t$ ) alors l'arc admet une tangente en tout point et le vecteur unitaire de la tangente est donné par

$$\tau(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|}, \quad (21.159)$$

pour tout  $t$  dans  $I$ .

CorUnitTgtaugnorma

**Corolaire 21.52.**

Si  $\gamma = (J, \gamma_N)$  est un arc paramétré de classe  $\mathcal{C}^1$ , normal, alors le vecteur unitaire de la tangente au point  $\gamma_N(s)$  est donné par  $\tau(s) = \gamma'_N(s)$ .

*Démonstration.* Nous devons démontrer que dans le cas d'un paramétrage normal nous avons  $\|\gamma'_N(s)\| = 1$  pour tout  $s$ . Par définition,

$$l([s, s'], g) = \int_s^{s'} \|\gamma'_N(u)\| du = s' - s. \quad (21.160)$$

Par conséquent,

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_s^{s+h} \|\gamma'_N(u)\| du = \lim_{h \rightarrow 0} \frac{s+h-s}{h} = 1. \quad (21.161)$$

Cela implique que  $\|\gamma'_N(s)\| = 1$ , et donc en particulier que  $(J, \gamma_N)$  est un arc régulier. Le corolaire précédent montre alors que  $\tau(s) = \gamma'_N(s)/\|\gamma'_N(s)\| = \gamma'_N(s)$ .  $\square$

**Exemple 21.53.**

Considérons la courbe  $\gamma(t) = (t^2, t^3)$ , et cherchons la tangente en  $t_0 = 0$ . En dérivant nous avons successivement

$$\begin{aligned} \gamma(t) &= (t^2, t^3) \\ \gamma'(t) &= (2t, 3t^2) \\ \gamma''(t) &= (2, 6t). \end{aligned} \quad (21.162)$$

En posant  $t = 0$ , nous trouvons que  $\gamma'(0) = 0$  mais  $\gamma''(0) = (2, 0) \neq 0$ . Le théorème nous dit donc que la direction de la tangente est horizontale. Nous pouvons faire le calcul directement :

$$\frac{\gamma(t) - \gamma(t_0)}{\|\gamma(t) - \gamma(t_0)\|} = \frac{(t^2, t^3)}{\sqrt{t^4 + t^6}} = \frac{(t^2, t^3)}{t^2 \sqrt{1 + t^2}} = \frac{(1, t)}{\sqrt{1 + t^2}}, \quad (21.163)$$

dont la limite  $t \rightarrow 0$  est bien le vecteur horizontal  $(1, 0)$ .

La figure 21.7 montre quelques tangentes, c'est-à-dire quelques vecteurs dans la direction  $\gamma'(t)$  (pour les  $t \neq 0$ , il ne faut pas aller à la dérivée seconde). Nous remarquons que de part et d'autres du sommet, les vecteurs ne sont pas dirigés dans le même sens. *En tant que vecteurs* de norme 1, ces vecteurs n'ont pas de limites quand  $t \rightarrow 0$ . Ce sont bien les *directions* qui ont une limite, parce que la direction ne tient pas compte du sens.

△

## 21.8 Un peu de topologie

La proposition 21.54 donne une sorte de théorème des valeurs intermédiaires pour le cas d'une application à valeurs dans un chemin.

PROPooJYGVooshNewy

**Proposition 21.54 ([1]).**

Soit une application continue et injective  $\gamma: [0, 1] \rightarrow \mathbb{R}^n$ . Nous posons  $\Gamma = \gamma([0, 1])$  sur lequel nous considérons la topologie induite<sup>11</sup> de  $\mathbb{R}^n$ .

11. Définition 7.27.

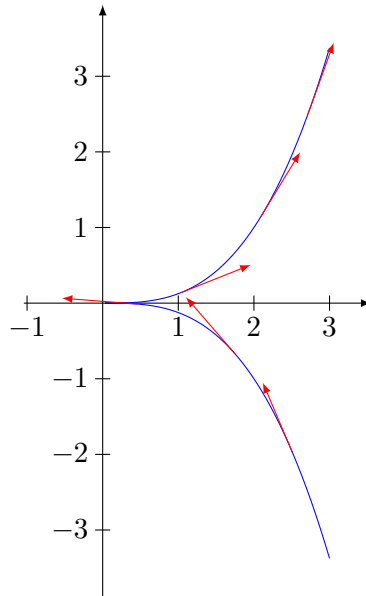


FIGURE 21.7: Quelques tangentes de la courbe  $\gamma(t) = (t^2, t^3)$ . LabelFigParamTangente

Nous supposons que  $\gamma^{-1}: \Gamma \rightarrow [0, 1]$  est continue<sup>12</sup>

Nous considérons un chemin  $\alpha: [0, 1] \rightarrow \mathbb{R}^n$  tel que

- (1)  $\alpha$  est continu,
- (2)  $\alpha(0) = \gamma(0)$
- (3)  $\alpha([0, 1]) \subset \Gamma$
- (4)  $\alpha(1) = \gamma(t_0)$  pour un certain  $t_0 \in [0, 1]$ .

Alors pour tout  $t \in [0, t_0]$ , il existe  $u \in [0, 1]$  tel que  $\alpha(u) = \gamma(t)$ .

*Démonstration.* Nous commençons par montrer que l'application  $\alpha: [0, 1] \rightarrow \Gamma$  est encore continue lorsque nous voyons bien l'espace d'arrivée comme  $\Gamma$  muni de sa propre topologie et non comme  $\mathbb{R}^n$  muni de sa topologie usuelle.

Soit un ouvert  $\mathcal{O}$  de  $\Gamma$ ; il existe un ouvert  $\mathcal{O}'$  de  $\mathbb{R}^n$  tel que  $\mathcal{O} = \mathcal{O}' \cap \Gamma$ . Vu que  $\alpha$  ne prend ses valeurs que dans  $\Gamma$ , nous avons  $\alpha^{-1}(\Gamma \cap \mathcal{O}') = \alpha^{-1}(\mathcal{O}')$  et comme  $\alpha$  est continue pour la topologie de  $\mathbb{R}^n$ , la partie  $\alpha^{-1}(\mathcal{O}')$  est un ouvert de  $[0, 1]$ .

Nous considérons donc l'application

$$\gamma^{-1} \circ \alpha: [0, 1] \rightarrow [0, 1] \quad (21.164)$$

qui est continue et vérifie donc le théorème des valeurs intermédiaires 9.95. Les hypothèses  $\alpha(0) = \gamma(0)$  et  $\alpha(1) = \gamma(t_0)$  donnent

$$(\gamma^{-1} \circ \alpha)(0) = 0 \quad (21.165a)$$

$$(\gamma^{-1} \circ \alpha)(1) = t_0. \quad (21.165b)$$

Donc pour tout  $t \in [0, t_0]$ , il existe  $u \in [0, 1]$  tel que  $(\gamma^{-1} \circ \alpha)(u) = t$ . En appliquant  $\gamma$  des deux côtés, nous voyons que ce  $u$  vérifie  $\alpha(u) = \gamma(t)$  comme demandé.  $\square$

La proposition suivante dit essentiellement que la longueur d'un chemin est minoré par la longueur de son graphe.

PROPooXENVooMvkTZW

**Proposition 21.55** ([1]).

Soient  $a, b \in \mathbb{R}^n$ . Soit une application continue et injective  $\gamma: [0, 1] \rightarrow \mathbb{R}^n$  telle que  $\gamma(0) = a$  et  $\gamma(1) = b$ .

12. Je ne suis pas certain que cette hypothèse soit indispensable.

Nous considérons maintenant un second chemin continu  $\alpha: [0, 1] \rightarrow \mathbb{R}^n$  tel que  $\alpha(0) = a$ ,  $\alpha(1) = b$  et  $\alpha([0, 1]) \subset \gamma([0, 1])$ .

Alors la longueur de  $\gamma$  minore celle de  $\alpha$  :  $l(\gamma) \leq l(\alpha)$ .

*Démonstration.* Pour rappel, la longueur d'un chemin est donné par le supremum des longueurs des lignes brisées reliant des points du chemin (définition 21.3). Soit une subdivision  $\sigma$  de  $[0, 1]$  et la « longueur » correspondante pour  $\gamma$  :

$$l_\sigma(\gamma) = \sum_{i=1}^n \|\gamma(t_i) - \gamma(t_{i-1})\|. \quad (21.166)$$

Nous allons construire une subdivision de  $[0, 1]$  pour laquelle la longueur de  $\alpha$  sera la même. De cette façon, le supremum pour  $\alpha$  sera un supremum sur un ensemble plus grand que celui du supremum pour  $\gamma$ .

Vu que le graphe de  $\alpha$  est contenu dans celui de  $\gamma$ , et que les points de départ et d'arrivée sont les mêmes, la proposition 21.54 nous donne que pour tout  $t \in [0, 1]$ , il existe  $u \in [0, 1]$  tel que  $\alpha(u) = \gamma(t)$ .

Cela ne suffit pas à considérer une subdivision  $(u_i)_{i=1, \dots, n}$  de  $[0, 1]$  pour laquelle  $\alpha(u_i) = \gamma(t_i)$  parce qu'il faut encore que les  $u_i$  soient ordonnés. Vu que  $\alpha$  n'est pas injective, il n'y a pas de garanties de ce côté. Nous posons :

$$u_i = \min\{x \in [0, 1] \text{ tel que } \alpha(x) = \gamma(t_i)\}. \quad (21.167)$$

Montrons que ce minimum existe. Ce minimum existe parce que  $\alpha$  étant continue, l'ensemble des  $x$  sur lesquels  $\alpha(x)$  a une valeur donnée est fermé (le complémentaire est ouvert par le théorème des valeurs intermédiaires). De plus cet ensemble est borné parce qu'il est inclus dans  $[0, 1]$ . Il est donc compact et possède un minimum.

Nous prouvons à présent que  $u_{i+1} \geq u_i$ . Nous avons  $\alpha(u_{i+1}) = \gamma(t_{i+1})$  et il existe  $u$  dans  $[0, u_{i+1}[$  tel que  $\alpha(u) = \gamma(t_i)$  (proposition 21.54). Vu que  $u_i$  est le minimum de tels  $u$ , nous avons  $u_i \in [0, u_{i+1}[$ . Cela prouve bien que  $u_i < u_{i+1}$ .

Ces  $(u_i)_{i=1, \dots, n}$  forment une subdivision de  $[0, 1]$  telle que  $\alpha(u_i) = \gamma(t_i)$ . La longueur associée à la subdivision  $(u_i)$  pour  $\alpha$  est la même que celle associée à  $(t_i)$  pour  $\gamma$ .

Donc  $l(\alpha) \geq l(\gamma)$ . □

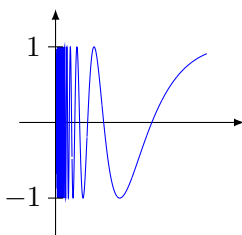


FIGURE 21.8: La figure de la proposition 21.56.

LabelFigYWxOakh

PROPooVXDNooPZYKPr

**Proposition 21.56** ([1]).

Nous considérons la partie suivante de  $\mathbb{R}^2$  :

$$A_2 = \{(t, \sin(1/t))\}_{t \in ]0, 1]} \quad (21.168)$$

qui est dessinée sur la figure 21.8. Ensuite nous posons

$$A = \{(0, 0)\} \cup A_2. \quad (21.169)$$

La partie  $A$  est connexe, mais pas connexe par arcs.

*Démonstration.* En plusieurs points.

- (1)  **$A_2$  est connexe** La partie  $A_2$  est l'image de la fonction

$$\begin{aligned} f: ]0, 1] &\rightarrow \mathbb{R}^2 \\ t &\mapsto (t, \sin(1/t)). \end{aligned} \quad (21.170)$$

Vu que  $f$  est continue et que son ensemble de départ est connexe,  $A_2$  est connexe (lemme 7.272).

- (2)  **$A$  est connexe** Soient deux ouverts disjoints  $\mathcal{O}_1$  et  $\mathcal{O}_2$  dont l'union contient  $A$ . Nous supposons que  $(0, 0) \in \mathcal{O}_1$ . Si  $\mathcal{O}_1$  contient  $B((0, 0), r)$ , alors il contient tous les points de la forme  $(\frac{1}{2k\pi}, 0)$  pour  $k$  assez grand. Ces points sont dans  $A_2$ .

Vu que  $\mathcal{O}_1$  contient des points de  $A_2$ , il doit contenir tous les points de  $A_2$ ; sinon les ouverts  $\mathcal{O}_1$  et  $\mathcal{O}_2$  contrediraient la connexité de  $A_2$ . Finalement,  $A \subset \mathcal{O}_1$  et  $A$  est connexe.

- (3) **Connexité par arcs : le chemin** Pour faire aller le chemin dans le bon sens, nous considérons

$$\begin{aligned} \gamma: [0, 1[ &\rightarrow \mathbb{R}^2 \\ t &\mapsto f(1-t). \end{aligned} \quad (21.171)$$

Si  $A$  était connexe par arcs, il existerait une application continue  $\alpha: [0, 1] \rightarrow A$  telle que  $\alpha(0) = (1, \sin(1)) = \gamma(0)$  et  $\alpha(1) = (0, 0)$ .

- (4) **Minoration de la longueur de  $\alpha$**  Vu que  $\alpha$  est continue, ses composantes le sont et vérifient le théorème des valeurs intermédiaires. Pour tout  $t \in ]0, 1[$ , il existe  $u \in ]0, 1[$  tel que  $\alpha(u) = \gamma(t)$ .

Par la proposition 21.55, la longueur de  $\alpha$  est minorée par la longueur du chemin  $\gamma$  entre 0 et  $t_0$  pour tout  $t$ .

- (5) **La longueur de  $\gamma$**  Vu que  $\gamma$  passe une infinité de fois par zéro et par 1, il est possible de construire une subdivision de  $]0, 1[$  par rapport à laquelle la longueur de  $\gamma$  est arbitrairement grande.

- (6) **Conclusion** Donc la longueur de  $\alpha$  est minorée par tous les nombres, arbitrairement grand. Autrement dit, la longueur de  $\alpha$  est infinie. Mais  $\alpha$  étant un chemin continu depuis un compact, cela est impossible.

Nous déduisons qu'un chemin continu liant  $(1, \sin(1))$  à  $(0, 0)$  en restant sur  $A$  est impossible. □

## 21.9 Repère de Frenet

Dans cette section, nous ne considérons que des courbes dans  $\mathbb{R}^3$ .

SecFrenet

### Proposition 21.57.

Soit  $\gamma = (J, \gamma_N)$  un arc paramétré normal de classe  $\mathcal{C}^2$ . Alors pour toute valeur de  $s$  dans  $J$ , nous avons

$$\tau(s) \cdot \tau'(s) = 0 \quad (21.172)$$

où  $\tau(s) = \gamma'_N(s)$ . C'est-à-dire que la dérivée seconde est perpendiculaire à la dérivée première.

*Démonstration.* Le paramétrage étant normal, nous avons

$$\|\gamma'_N(s)\|^2 = \sum_{i=1}^n x'_i(s)^2 = 1; \quad (21.173)$$

ce qui implique, en dérivant les deux membres, que

$$0 = 2 \sum_{i=1}^n x'_i(s) x''_i(s), \quad (21.174)$$

c'est-à-dire exactement  $\gamma'_N(s) \cdot \gamma''_N(s) = 0$ ; d'où la thèse. □

Proptausclataupzero

**Remarque 21.58.**

Si nous n'utilisons pas des coordonnées normales, la proposition 21.57 n'est pas spécialement vraie. Prenons par exemple la courbe qui donne la parabole :

$$\gamma(t) = (t, t^2) \quad (21.175a)$$

$$\gamma'(t) = (1, 2t) \quad (21.175b)$$

$$\gamma''(t) = (0, 2) \quad (21.175c)$$

Nous avons  $\gamma'(t) \cdot \gamma''(t) = 4t$ . Par conséquent, la dérivée seconde n'est la normale à la courbe que en  $t = 0$ . Cela est une propriété très intéressante des coordonnées normales : la dérivée seconde d'une coordonnées normale donne un vecteur normal à la courbe, c'est-à-dire perpendiculaire à la tangente.

DefCourbureNormleUnit

**Définition 21.59.**

Soit  $\gamma = (J, \gamma_N)$  un arc paramétré normal de classe  $\mathcal{C}^2$ .

(1) Le **vecteur unitaire tangent** est donné par le corolaire 21.51 :  $\tau(t) = \gamma'(t)/\|\gamma'(t)\|$ .

(2) La **normale principale** est le vecteur  $\tau'(s)$ . Le **vecteur unitaire normal** est le vecteur

$$\nu(s) = \frac{\tau'(s)}{\|\tau'(s)\|} = \frac{\gamma_N''(s)}{\|\gamma_N''(s)\|}. \quad (21.176)$$

Nous déduirons une formule plus pratique en dehors des coordonnées normales en (21.204).

(3) La **courbure** au point  $\gamma_N(s)$  est le réel

$$c(s) = \|\tau'(s)\| = \|\gamma_N''(s)\|. \quad (21.177)$$

Note : il y a une notion de courbure signée qui sera donnée dans la définition 21.73.

(4) Le **rayon de courbure** est le réel

$$R(s) = \frac{1}{c(s)} = \frac{1}{\|\gamma_N''(s)\|}. \quad (21.178)$$

Par la proposition 21.57, nous avons  $\nu(s) \cdot \tau(s) = 0$ . En combinant toutes les formules, nous avons les différentes expressions suivantes pour le vecteur normal unitaire :

$$\nu(s) = \frac{\gamma_N''(s)}{c(s)} = \frac{\tau'(s)}{\|\tau'(s)\|} = \frac{\tau'(s)}{c(s)} = R(s)\tau'(s) = R(s)\gamma_N''(s). \quad (21.179)$$

Eq0908nufractauRc

**Proposition 21.60.**

La fonction courbure s'écrit  $c = \|\gamma_N' \times \gamma_N''\|$ .

*Démonstration.* Par la proposition 10.43 nous avons :

$$\langle \gamma_N', \gamma_N'' \rangle^2 + \|\gamma_N' \times \gamma_N''\|^2 = \|\gamma_N'\|^2 \|\gamma_N''\|^2 = \|\gamma_N''\|^2 \quad (21.180)$$

parce que, le paramétrage étant normal,  $\|\gamma_N'\| = 1$ . Mais  $\langle \gamma_N', \gamma_N'' \rangle = 0$ , donc il reste  $\|\gamma_N' \times \gamma_N''\|^2 = \|\gamma_N''\|^2$ , d'où

$$c(s) = \|\gamma_N''(s)\| = \|\gamma_N'(s) \times \gamma_N''(s)\| \quad (21.181)$$

Eqcsnormmgppps

pour chaque  $s$  dans  $J$ . □

**Définition 21.61.**

Soit  $s$  un point birégulier (c'est-à-dire  $\gamma_N'(s) \neq 0$  et  $\gamma_N''(s) \neq 0$ ) de l'arc normal  $\gamma = (J, \gamma_N)$ . Le **vecteur unitaire de la binormale** est le vecteur

$$\beta(s) = \tau(s) \times \nu(s) \quad (21.182)$$

Par leurs définitions,  $\tau$  et  $\nu$  sont unitaires, tandis que la proposition 21.57 montre qu'ils sont également orthogonaux. Les propriétés du produit vectoriel font que  $\beta$  est également unitaire, et simultanément orthogonal à  $\tau$  et à  $\nu$ .

**Définition 21.62.**

Le repère orthonormal  $\{\tau(s), \nu(s), \beta(s)\}$  est le **repère de Frenet** au point  $\gamma_N(s)$ .

**Lemme 21.63.**

Le vecteur unitaire normal est donné par  $\nu(s) = \beta(s) \times \tau(s)$ .

*Démonstration.* Ceci est une application de la formule d'expulsion (10.121) et de l'orthonormalité de la base de Frenet :

$$\beta \times \tau = (\tau \times \nu) \times \tau = -(\nu \cdot \tau)\tau + (\tau \cdot \tau)\nu = \nu. \quad (21.183)$$

□

### 21.9.1 Torsion

Décomposons le vecteur  $\beta'(s)$  dans la base de Frenet. Pour cela nous allons utiliser la proposition 8.170 et montrer que  $\beta'(s) \cdot \tau(s) = \beta'(s) \cdot \beta(s) = 0$ , ce qui voudra dire que, dans la base de Frenet, les composantes de  $\beta'$  le long de  $\tau$  et  $\beta$  sont nulles. Le vecteur  $\beta'$  sera donc colinéaire à  $\nu$ .

D'abord, étant donné que la norme de  $\beta(s)$  est constante par rapport à  $s$ , nous avons

$$0 = \frac{d}{ds} \|\beta(s)\|^2 = 2\beta'(s) \cdot \beta(s). \quad (21.184)$$

Ensuite, nous dérivons la définition  $\beta(s) = \tau(s) \times \nu(s)$  en utilisant la formule de Leibniz (12.490) :

$$\beta'(s) = \tau'(s) \times \nu(s) + \tau(s) \times \nu'(s). \quad (21.185)$$

Mais  $\tau'(s) = \gamma_N''(s)$  tandis que  $\nu(s) = \frac{\gamma_N''(s)}{\|\gamma_N''(s)\|}$ , de telle sorte que  $\tau'(s) \times \nu(s) = 0$ . Nous restons donc avec  $\beta'(s) = \tau(s) \times \nu'(s)$ , ce qui prouve que  $\beta'(s)$  est perpendiculaire à  $\tau(s)$  et donc que  $\beta'(s) \cdot \tau(s) = 0$ .

Le vecteur  $\beta'(s)$  est donc un multiple de  $\nu(s)$ . Nous notons  $t(s)$  le facteur de proportionnalité :

$$\beta'(s) = t(s)\nu(s). \quad (21.186)$$

DefTorsion

**Définition 21.64.**

Soit  $\gamma = (J, \gamma_N)$  un arc paramétré normal de classe  $\mathcal{C}^3$ . La **torsion** de  $\gamma$  au point  $\gamma_N(s)$  est le réel

$$t(s) = \|\beta'(s)\| = \|\tau(s) \times \nu'(s)\|. \quad (21.187)$$

Lorsque  $t(s) \neq 0$ , le réel  $T(s) = \frac{1}{t(s)}$  est le **rayon de torsion** de  $\gamma$  en  $\gamma_N(s)$ .

Étant donné que pour chaque  $s$ , l'ensemble  $\{\tau(s), \nu(s), \beta(s)\}$  est une base, il est naturel de vouloir décomposer leurs dérivées dans cette base. D'abord, par définition de  $c$  et de  $t$ , nous avons

$$\begin{aligned} \tau'(s) &= c(s)\nu(s) \\ \beta'(s) &= t(s)\nu(s). \end{aligned} \quad (21.188)$$

Il reste à décomposer  $\nu'(s)$ . Définissons  $\alpha_\tau$ ,  $\alpha_\nu$  et  $\alpha_\beta$  (qui peuvent dépendre de  $s$ ) par

$$\nu'(s) = \alpha_\tau \tau(s) + \alpha_\nu \nu(s) + \alpha_\beta \beta(s). \quad (21.189)$$

En vertu de la proposition 8.170, nous avons

$$\begin{aligned} \alpha_\tau &= \langle \nu'(s), \tau(s) \rangle = -\langle \nu(s), \tau'(s) \rangle = -\langle \nu(s), c(s)\nu(s) \rangle = -c(s), \\ \alpha_\nu &= \langle \nu'(s), \nu(s) \rangle = 0, \\ \alpha_\beta &= \langle \nu'(s), \beta(s) \rangle = -\langle \nu(s), \beta'(s) \rangle = -t(s), \end{aligned} \quad (21.190)$$

où nous avons utilisé le fait que  $\langle \nu(s), \nu(s) \rangle = \|\nu(s)\|^2 = 1$ . Si nous mettons ces résultats sous forme matricielle, nous avons les **formules de Frenet** :

$$\begin{pmatrix} \tau'(s) \\ \nu'(s) \\ \beta'(s) \end{pmatrix} = \begin{pmatrix} 0 & c(s) & 0 \\ -c(s) & 0 & -t(s) \\ 0 & t(s) & 0 \end{pmatrix} \begin{pmatrix} \tau(s) \\ \nu(s) \\ \beta(s) \end{pmatrix}. \quad (21.191)$$

**Proposition 21.65.**

Si  $s$  est un point birégulier, alors la torsion est donnée par

$$t(s) = -\frac{(\gamma'_N \times \gamma''_N) \times \gamma'''_N}{\|\gamma'_N(s) \times \gamma''_N(s)\|^2}. \quad (21.192)$$

*Démonstration.* Par l'équation (21.179), nous avons  $\gamma''_N(s) = c(s)\nu(s)$ , et par conséquent

$$\gamma'''_N(s) = c'(s)\nu(s) + c(s)\nu'(s) = c'(s)\nu(s) + c(s)[-c(s)\tau(s) - t(s)\beta(s)], \quad (21.193)$$

où nous avons utilisé la formule de Frenet pour  $\nu'(s)$ . Par ailleurs, sachant le corolaire 21.52 et la formule de Frenet pour  $\tau'$ , nous avons

$$\gamma'_N \times \gamma''_N = \tau(s) \times \tau'(s) = \tau(s) \times c(s)\nu(s) = c(s)\beta(s). \quad (21.194)$$

En combinant les deux dernières équations, et en se souvenant que la base de Frenet est orthonormale,

$$(\gamma'_N \times \gamma''_N) \cdot \gamma'''_N(s) = -c(s)^2 t(s), \quad (21.195)$$

et donc, en remplaçant  $c(s)$  par la formule (21.181),

$$t(s) = -\frac{(\gamma'_N \times \gamma''_N) \cdot \gamma'''_N}{\|\gamma'_N \times \gamma''_N\|^2}. \quad (21.196)$$

□

## 21.10 Hors des coordonnées normales

Remfougnormoupad

**Remarque 21.66.**

Notons que la définition de  $\tau$  est donnée pour tout arc  $\mathcal{C}^1$  régulier  $(I, \gamma)$  par  $\tau(t) = \gamma'(t)/\|\gamma'(t)\|$ . La propriété  $\tau = \gamma'_N$  n'est valable que lorsque le paramétrage est normal. Les autres définitions ont toutes été données dans le cas d'un paramétrage normal.

La remarque 21.66 nous incite à exprimer toute la base de Frenet en termes de  $\gamma$  lorsque le paramétrage n'est pas normal. Étant donné que nous pouvons toujours faire le changement de variable  $\gamma(t) = \gamma_N(\phi(t))$  (proposition 21.46), il est possible d'exprimer les vecteurs  $\tau$ ,  $\nu$  et  $\beta$  ainsi que les réels  $c$  et  $t$  en fonction de  $\gamma$  et de ses dérivées.

Nous allons maintenant travailler à écrire les formules.

Pour plus de facilité, nous collectons les définitions. Afin d'alléger la notation, nous n'exprimons pas explicitement les dépendances en  $s$  :

**Vecteur unitaire tangent** Par le corolaire 21.52,  $\tau$  est donné par  $\tau = \gamma'_N$ .

**Vecteur unitaire normal** Par la définition 21.59,  $\nu$  est donné par  $\nu = \frac{\tau'}{\|\tau'\|}$ .

**Vecteur unitaire de la binormale** Par la définition 21.59,  $\beta$  est donné par  $\beta = \tau \times \nu$ .

**Courbure** Par la définition 21.59,  $c$  est donné par  $c = \|\tau'\|$ .

**Torsion** Par la définition 21.64,  $t$  est donné par  $t = \|\beta'\|$ .



Le schéma du changement de variable est

$$\begin{array}{ccc} t \in I & \xrightarrow{\gamma} & \mathbb{R}^3 \\ \phi \downarrow & \nearrow \gamma_N & \\ s \in J & & \end{array} \quad \text{EqDiagJstgvpR} \quad (21.197)$$

La difficulté ne sera pas d'éliminer  $\gamma_N$  de toutes les formules, mais bien de se débarrasser des fonctions  $\phi$  qui arrivent quand nous exprimons  $\gamma_N$  en termes de  $\gamma$ , et en particulier lorsque nous voulons exprimer les dérivées de  $\gamma_N$  en termes de  $\gamma$  et de ses dérivées.

Regardons d'abord comment les dérivées de  $\gamma_N$  s'expriment en termes de  $\gamma$ . En utilisant le fait que  $\gamma_N(s) = (\gamma \circ \phi^{-1})(s)$  et que  $\|\gamma'_N(s)\| = 1$ , nous avons

$$\gamma'_N(s) = \frac{\gamma'_N(s)}{\|\gamma'_N(s)\|} = \frac{(\gamma \circ \phi^{-1})'(s)}{\|(\gamma \circ \phi^{-1})'(s)\|} = \frac{\gamma'(\phi^{-1}(s))(\phi^{-1})'(s)}{\|\gamma'(\phi^{-1}(s))\| |(\phi^{-1})'(s)|} = \frac{\gamma'(t)}{\|\gamma'(t)\|} \quad \text{EqgNgppNur} \quad (21.198)$$

où nous avons utilisé le fait que  $\phi^{-1}$  étant croissante (parce que l'inverse d'une fonction croissante est croissante),  $(\phi^{-1})'(s) = |(\phi^{-1})'(s)|$ . Pourquoi écrivons nous  $|(\phi^{-1})'(s)|$  et non  $\|\phi^{-1}(s)\|$  ?

Pour la dérivée seconde, nous dérivons la relation (21.198) :

$$\gamma''_N(s) = \frac{\gamma''(\phi^{-1}(s))(\phi^{-1})'(s)}{\|\gamma'(\phi^{-1}(s))\|} + \gamma'(\phi^{-1}(s)) \frac{d}{ds} \left[ \frac{1}{\|\gamma'(\phi^{-1}(s))\|} \right]. \quad (21.199)$$

Le petit calcul suivant va nous permettre de simplifier cette expression :

$$(\phi^{-1})'(s) = (\phi^{-1})'(\phi(t)) = \frac{1}{\phi'(t)} = \frac{1}{\|\gamma'(t)\|}. \quad \text{Eavpmetfppnm} \quad (21.200)$$

Donc

$$\gamma''_N(s) = \frac{\gamma''(t)}{\|\gamma'(t)\|^2} + \gamma'(t) \frac{d}{ds} \left[ \frac{1}{\|\gamma'(t)\|} \right] \quad (21.201)$$

où il est entendu que  $t = \phi^{-1}(s)$ . Avec cette expression, nous ne nous sommes pas encore débarrassés de la fonction  $\phi$ , mais nous allons voir que cela nous sera suffisant.

Pour le vecteur unitaire tangent  $\tau(s)$ , nous avons donc immédiatement

$$\tau(s) = \gamma'_N(s) = \frac{\gamma'(t)}{\|\gamma'(t)\|}. \quad \text{EqTauvect} \quad (21.202)$$

Ici encore il est sous-entendu que le  $t$  dans le membre de droite est lié au  $s$  du membre de gauche par  $t = \phi^{-1}(s)$ . Il est donc naturel de nous demander si nous avons gagné quelque chose, étant donné que la formule (21.202) contient encore la fonction  $\phi$ .

Géométriquement, le vecteur  $\tau(s)$  est le vecteur normal unitaire de la courbe au point  $\gamma_N(s)$ . En utilisant les relations du diagramme (21.197), nous avons en réalité  $\gamma_N(s) = \gamma_N(\phi(t)) = \gamma(t)$ .

Le vecteur  $\frac{\gamma'(t)}{\|\gamma'(t)\|}$  représente donc le vecteur normal tangent au point  $\gamma(t)$ .

Pour calculer la courbure, nous devons d'abord calculer le produit vectoriel

$$\begin{aligned} \gamma'_N(s) \times \gamma''_N(s) &= \frac{\gamma'(t)}{\|\gamma'(t)\|} \times \left( \frac{\gamma''(t)}{\|\gamma'(t)\|^2} + \gamma'(t) \frac{d}{ds} \left[ \frac{1}{\|\gamma'(t)\|} \right] \right) \\ &= \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t)\|^3} \end{aligned} \quad \text{eqProductogpp} \quad (21.203)$$

parce que le deuxième terme dans la parenthèse est un multiple de  $\gamma'(t)$ , de telle sorte que son produit vectoriel avec  $\gamma'(t)/\|\gamma'(t)\|$  soit nul. En prenant la norme,

$$c(s) = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3}. \quad \text{EqCourburetermf} \quad (21.204)$$

Encore une fois, cette équation nous enseigne que la courbure au point  $\gamma(t) \in \mathbb{R}^3$  est donnée par le membre de droite, qui ne dépend que de  $t$ .

Le vecteur unitaire binormal est donné par  $\beta(s) = \tau(s) \times \nu(s)$ . En utilisant (21.202) et (21.179),

$$\beta(s) = \tau(s) \times \nu(s) = \gamma'_N(s) \times \frac{\gamma''_N(s)}{c(s)}. \quad (21.205)$$

Les formules (21.203) pour le produit vectoriel et (21.204) pour la courbure donnent ensuite

$$\beta(s) = \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t)\|^3} \cdot \frac{1}{c(s)} = \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t) \times \gamma''(t)\|}. \quad (21.206)$$

Cela donne le vecteur unitaire binormal au point  $\gamma(t)$  en termes de  $\gamma'(t)$  et  $\gamma''(t)$ .

La torsion demande d'utiliser la dérivée troisième de  $\gamma_N$ . Nous avons

$$\begin{aligned} \gamma'''_N(s) &= (\gamma \circ \phi^{-1})'''(s) \\ &= \left( \gamma'(\phi^{-1}(s))(\phi^{-1})'(s) \right)'' \\ &= \left( \gamma''(\phi^{-1}(s))(\phi^{-1})'(s)^2 + \gamma'(\phi^{-1}(s))(\phi^{-1})''(s) \right)' \\ &= \gamma'''(\phi^{-1}(s))(\phi^{-1})'(s)^3 + v \\ &= \frac{\gamma'''(\phi^{-1}(s))}{\|\gamma'(t)\|^3} + v \end{aligned} \quad \text{par (21.200)} \quad (21.207)$$

où  $v$  est un élément de  $\text{Span}\{\gamma''(\phi^{-1}(s)), \gamma'(\phi^{-1}(s))\}$ . Le vecteur  $v$  est donc perpendiculaire à  $\gamma' \times \gamma''$  et donc à  $\gamma'_N \times \gamma''_N$  à cause de la relation (21.203) qui montre que  $\gamma' \times \gamma''$  est parallèle à  $\gamma'_N \times \gamma''_N$ . De ce fait, lorsque nous calculons  $(\gamma'_N \times \gamma''_N) \cdot \gamma'''_N$ , la partie  $v$  de  $\gamma'''_N$  n'entre pas en ligne de compte.

Nous avons donc le calcul suivant, en remplaçant les diverses occurrences de  $\gamma'_N \times \gamma''_N$  par sa valeur (21.203) en termes de  $\gamma$ ,

$$\begin{aligned} t(s) &= -\frac{(\gamma'_N \times \gamma''_N) \cdot \gamma'''_N}{\|\gamma'_N \times \gamma''_N\|^2} \\ &= -\frac{(\gamma'_N \times \gamma''_N) \cdot \gamma'''(t)}{\|\gamma'_N \times \gamma''_N\|^2 \|\gamma'(t)\|^2} \\ &= -\frac{(\gamma' \times \gamma'') \cdot \gamma'''}{\|\gamma' \times \gamma''\|^2}. \end{aligned} \quad (21.208)$$

Dans cette expression, il est sous-entendu que tous les  $\gamma_N$  sont fonctions de  $s$  et tous les  $\gamma$  sont fonction de  $t$  où  $s$  et  $t$  sont liés par  $s = \phi(t)$ .

Voici un théorème qui résume ce que nous venons de faire.

THOooXOVRooRXtEuU

### **Théorème 21.67.**

Pour tout représentant  $(I, \gamma)$ , les éléments métriques  $(\tau, \nu, \beta, c, t)$  au point  $\gamma(t)$  s'expriment en fonction de  $\gamma(t)$ ,  $\gamma'(t)$ ,  $\gamma''(t)$  et  $\gamma'''(t)$ .

### **Lemme 21.68.**

Si  $\gamma$  est le graphe de la fonction  $y$  alors la courbure de  $\gamma$  est donnée par la formule

$$c(\gamma(t)) = \frac{|y''(t)|}{(1 + y'(t)^2)^{3/2}} \quad (21.209)$$

*Démonstration.* Nous avons :

$$\gamma(t) = (t, y(t)) \quad (21.210a)$$

$$\gamma'(t) = (1, y'(t)) \quad (21.210b)$$

$$\gamma''(t) = (0, y''(t)). \quad (21.210c)$$

Il s'agit maintenant seulement d'utiliser la formule (21.204) en se souvenant comment on calcule un produit vectoriel <sup>13</sup> :

$$\gamma' \times \gamma'' = \begin{vmatrix} e_1 & e_2 & e_3 \\ 1 & y' & 0 \\ 0 & y'' & 0 \end{vmatrix} = y'' e_1. \quad (21.211)$$

□

## 21.11 Tracer des courbes paramétriques dans $\mathbb{R}^2$

SecTracerParmCourbe

Nous allons maintenant voir comment les concepts introduits nous aident à effectivement tracer des courbes dans le plan. Les courbes que nous regardons sont de la forme  $\gamma(t) = (x(t), y(t))$ , et nous supposons que ces fonctions soient suffisamment régulières (disons trois fois continument dérivables). Nous ne supposons pas que la courbe soit donnée en coordonnées normales, en particulier,  $\gamma''(t)$  n'est pas le vecteur normal en  $\gamma(t)$ .

La notion clef qui va jouer est le **cercle osculateur** de la courbe  $\gamma$  au point  $\gamma(t)$ . Sans rentrer dans les détails, disons que c'est le cercle qui « colle » le mieux possible la courbe. Le rayon de ce cercle est le rayon de courbure :

$$R(t) = \frac{\|\gamma(t)\|^3}{\|\gamma'(t) \times \gamma''(t)\|}. \quad (21.212)$$

En pratique, le produit vectoriel se calcule comme ceci :

$$\gamma'(t) \times \gamma''(t) = \begin{vmatrix} e_x & e_y & e_z \\ x'(t) & y'(t) & 0 \\ x''(t) & y''(t) & 0 \end{vmatrix} = (x'y'' - x''y')e_z. \quad (21.213)$$

Le centre du cercle osculateur va se trouver quelque part sur la normale. Le vecteur normal est donné par

$$n(t) = J \frac{\gamma'(t)}{\|\gamma'(t)\|} \quad (21.214)$$

où  $J$  est la rotation d'angle  $\frac{\pi}{2}$  :

$$J \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} y'(t) \\ -x'(t) \end{pmatrix}. \quad (21.215)$$

Cela nous laisse deux possibilités pour le centre du cercle osculateur :  $\gamma(t) + R(t)n(t)$  ou bien  $\gamma(t) - R(t)n(t)$ . Il faut savoir de quel côté de la courbe est situé le centre du cercle osculateur. Il faut choisir le côté de la concavité, c'est-à-dire le côté de la dérivée seconde.

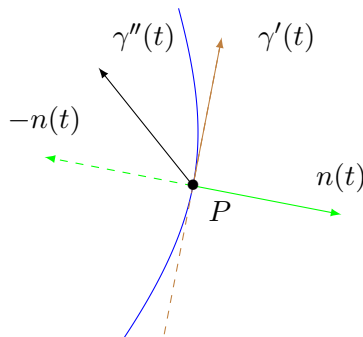


FIGURE 21.9: De quel côté de  $\gamma'(t)$  se trouvent  $n(t)$  et  $-n(t)$  ?

LabelFigQuelCote

La difficulté maintenant est de savoir si  $n(t)$  ou  $-n(t)$  est du côté de  $\gamma''(t)$ . Il faut savoir si  $n(t)$  est du même côté de la droite tangente que  $\gamma''(t)$  ou non. Par construction, si nous regardons

13. Définition 10.36.

la figure 21.9, le vecteur  $n(t)$  sera toujours à gauche de  $\gamma'(t)$ . Le fait que  $\gamma''(t)$  soit à gauche ou à droite de  $\gamma'(t)$  est donné par le signe du produit vectoriel  $\gamma'(t) \times \gamma''(t)$ . Si ce produit vectoriel est positif, il faut choisir  $-n(t)$  et si il est négatif, il faut choisir  $n'(t)$ .

Le truc pour obtenir le signe de  $x'y'' - x''y'$  est de faire

$$\frac{(\gamma' \times \gamma'') \cdot e_z}{\|\gamma' \times \gamma''\|}. \quad (21.216)$$

Le centre de courbure sera donc situé à la position

$$\Omega(t) = \gamma(t) - n(t) \frac{\|\gamma(t)\|^3}{\|\gamma'(t) \times \gamma''(t)\|^2} (\gamma' \times \gamma'') \cdot e_z \quad (21.217)$$

Nous pouvons écrire cela plus explicitement en nous souvenant que  $\gamma' \times \gamma'' = (x'y'' - x''y')e_z$ , par conséquent  $\frac{(\gamma' \times \gamma'') \cdot e_z}{\|\gamma' \times \gamma''\|^2} = \frac{1}{x'y'' - x''y'}$ . Nous avons

$$\Omega_x(t) = x(t) - y'(t) \frac{x'^2 + y'^2}{x'y'' - x''y'} \quad (21.218a)$$

$$\Omega_y(t) = y(t) + x'(t) \frac{x'^2 + y'^2}{x'y'' - x''y'}. \quad (21.218b)$$

Quelques exemples de cercles osculateurs sont sur la figure 21.10.

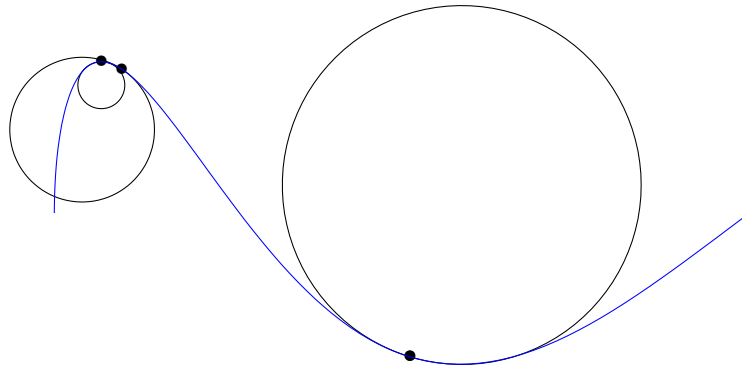


FIGURE 21.10: Exemple de cercles osculateurs.

LabelFigOsculateur

## 21.12 Courbes planes

### Définition 21.69.

Une courbe  $\gamma: [a, b] \rightarrow \mathbb{R}^n$  est **fermée** si  $\gamma(a) = \gamma(b)$ . Elle est **simple** si  $\gamma(t) \neq \gamma(t')$  dès que  $t, t' \in ]a, b[$  et  $t \neq t'$ .

DEFooSAZTooZGQrQG

### Définition 21.70.

Nous disons qu'une courbe fermée est continue, de classe  $C^1$ , de classe  $C^2$  ou autre condition de régularité si son extension périodique comme application  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$  a cette régularité.

### 21.12.1 Angle

LEMooUECMooNBDGiR

#### Lemme 21.71 ([603]).

Soient des courbes régulières  $\gamma$  et  $\sigma$  de classe  $C^2$  de l'intervalle ouvert  $I$  vers  $\mathbb{R}^2$ . Soit  $\theta_0 \in \mathbb{R}$  tel que

$$\frac{\gamma'(t_0) \cdot \sigma'(t_0)}{\|\gamma'(t_0)\| \|\sigma'(t_0)\|} = \cos(\theta_0) \quad (21.219a)$$

$$\frac{\gamma'(t_0) \cdot J\sigma'(t_0)}{\|\gamma'(t_0)\| \|\sigma'(t_0)\|} = \sin(\theta_0). \quad (21.219b)$$

Alors il existe une unique fonction différentiable  $\theta: I \rightarrow \mathbb{R}$

$$\frac{\gamma'(t) \cdot \sigma'(t)}{\|\gamma'(t)\| \|\sigma'(t)\|} = \cos(\theta(t)) \quad (21.220a)$$

$$\frac{\gamma'(t) \cdot J\sigma'(t)}{\|\gamma'(t)\| \|\sigma'(t)\|} = \sin(\theta(t)). \quad (21.220b)$$

*Démonstration.* Il suffit de prendre

$$f(t) = \frac{\gamma'(t) \cdot \sigma'(t)}{\|\gamma'(t)\| \|\sigma'(t)\|} \quad (21.221)$$

et

$$g(t) = \frac{\gamma'(t) \cdot J\sigma'(t)}{\|\gamma'(t)\| \|\sigma'(t)\|} \quad (21.222)$$

dans la proposition 18.58. Ces courbes sont de classe  $C^1$  parce que  $\gamma$  et  $\sigma$  sont de classe  $C^2$ .  $\square$

### 21.12.2 Courbure signée

Nous avons déjà défini la courbure d'une courbe en la définition 21.59. Nous introduisons maintenant la courbure signée qui est propre à la dimension deux.

DEFooTSJXooTlyRXf

#### Définition 21.72.

La **structure complexe** sur  $\mathbb{R}^2$  est l'application

$$\begin{aligned} J: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (x, y) &\mapsto (-y, x). \end{aligned} \quad (21.223)$$

DEFooJFWEEooXcIVUs

#### Définition 21.73.

La **courbure signée** de la courbe  $\gamma: I \rightarrow \mathbb{R}^2$  ( $I$  est un intervalle dans  $\mathbb{R}$ ) est la fonction

$$\kappa(t) = \frac{\gamma''(t) \cdot J\gamma'(t)}{\|\gamma'(t)\|^3} \quad (21.224)$$

EQooWOUQooXrVzGx

où  $J$  est la structure complexe de la définition 21.72.

Cette définition est motivée par le fait qu'en identifiant  $\mathbb{R}^2$  à  $\mathbb{C}$ , l'application  $J$  revient à l'application  $z \mapsto iz$ .

Si  $v, w \in \mathbb{R}^2$  nous avons formellement

$$v \times w = -(v \cdot Jw)e_3. \quad (21.225)$$

En particulier pour tout  $v \in \mathbb{R}^2$  nous avons

$$v \cdot Jv = 0. \quad (21.226)$$

#### Lemme 21.74.

Soit une courbe régulière  $\gamma: [a, b] \rightarrow \mathbb{R}^2$  et un difféomorphisme  $h: [c, d] \rightarrow [a, b]$ . Si nous posons  $\sigma = \gamma \circ h$  alors

$$\kappa_\sigma(u) = \operatorname{sgn}(h'(u))\kappa_\gamma(h(u)). \quad (21.227)$$

EQooSQNMooUKGhPd

*Démonstration.* Nous utilisons la définition (21.224) de la courbure signée. La règle de dérivation en chaîne donne :

$$\sigma'(u) = \gamma'(h(u))h'(u) \quad (21.228a)$$

$$\sigma''(u) = \gamma''(h(u))h'(u)^2 + \gamma'(h(u))h''(u). \quad (21.228b)$$

La numérateur de  $\kappa_\sigma(u)$  est :

$$(\gamma'' \circ h)h'^2 \cdot J(\gamma' \circ h)h' + h''(\gamma' \circ h) \cdot J(\gamma' \cdot h)h' \quad (21.229)$$

dont le second terme est nul parce que  $v \cdot Jv = 0$ . Il nous reste donc

$$\kappa_\sigma(u) = \frac{(h')^3}{|h'|^3} \frac{(\gamma'' \circ h) \cdot J(\gamma' \circ h)}{\|\gamma' \circ h\|^3} = \operatorname{sgn}(h') \kappa_\gamma(h(u)). \quad (21.230)$$

□

LEM00KFOR00EGJCrm

**Lemme 21.75** ([603]).

Si  $\gamma_N$  est un arc paramétré normal, alors

$$\gamma_N''(s) = \kappa(s) J\gamma_N'(s). \quad (21.231)$$

*Démonstration.* Vu que le paramétrage est normal,  $\gamma_N' \cdot \gamma_N' = 1$ , et en dérivant,  $\gamma_N'' \cdot \gamma_N' = 0$ . Donc  $\gamma_N''$  est un multiple de  $J\gamma_N'$ . En tenant compte du fait que le paramétrage est normal, la courbure est

$$\kappa(s) = \gamma_N''(s) \cdot J\gamma_N'(s). \quad (21.232)$$

En y injectant  $\gamma_N''(s) = \lambda(s) J\gamma_N'(s)$  nous trouvons

$$\kappa(s) = \lambda(s) J\gamma_N'(s) \cdot J\gamma_N'(s) = \lambda(s). \quad (21.233)$$

Donc le facteur de proportionnalité est  $\kappa(s)$ .

□

THO00DLDV00FQnLWn

**Théorème 21.76.**

Soit une courbe  $\gamma: I \rightarrow \mathbb{R}^2$  de classe  $C^2$ .

(1)  $\gamma$  est une partie de droite si et seulement si  $\kappa(t) = 0$  pour tout  $t$ .

(2)  $\gamma$  est une partie d'un cercle de rayon  $r > 0$  si et seulement si  $|\kappa(s)| = \frac{1}{r}$ .

*Démonstration.* Si  $\gamma$  est une droite, la dérivée seconde est nulle et la courbure est nulle. Supposons pour la réciproque que  $\kappa(t) = 0$  pour tout  $t$ . Nous utilisons un paramétrage normal de  $\gamma$ , ce qui ne change pas que la courbure reste nulle. Nous avons par le lemme 21.75 que  $\gamma''(s) = 0$  et donc l'existence de  $a, b \in \mathbb{R}^2$  tels que  $\gamma(t) = at + b$ .

Passons au cas du cercle. Si  $\gamma$  est un cercle, le paramétrage normal est

$$\gamma(t) = R \begin{pmatrix} \cos(\frac{t}{R}) \\ \sin(\frac{t}{R}) \end{pmatrix} \quad (21.234a)$$

$$\gamma'(t) = \begin{pmatrix} -\sin(\frac{t}{R}) \\ \cos(\frac{t}{R}) \end{pmatrix} \quad (21.234b)$$

$$\gamma''(t) = -\frac{1}{R} \begin{pmatrix} \cos(\frac{t}{R}) \\ \sin(\frac{t}{R}) \end{pmatrix}. \quad (21.234c)$$

Avec tout cela nous avons  $\kappa(s) = \gamma''(s) \cdot J\gamma'(s) = \frac{1}{R}$ .

Nous supposons enfin que  $\kappa(t) = 1/R$  et que le paramétrage soit normal (encore une fois, un reparamétrage ne change pas la courbure lorsqu'elle est constante). Nous définissons la courbe

$$\begin{aligned} \beta: [b, c] &\rightarrow \mathbb{R}^2 \\ t &\mapsto \gamma(t) + RJ\gamma'(t). \end{aligned} \quad (21.235)$$

Nous avons  $\beta'(t) = \gamma'(t) + RJ\gamma''(t)$ . Mais par le lemme 21.75 nous avons  $\gamma'' = kJ\gamma' = \frac{1}{R}J\gamma'$ . Donc

$$\beta'(t) = \gamma'(t) - \gamma'(t) = 0. \quad (21.236)$$

Du coup  $\beta$  est constante :  $\beta(t) = a$ . Alors  $a = \gamma(t) + RJ\gamma'(t)$  et en particulier

$$\|\gamma(t) - a\| = \|RJ\gamma'(t)\| = r. \quad (21.237)$$

Donc effectivement  $\gamma$  reste sur un cercle de rayon  $r$  et de centre  $a$ .

□

DEFooFAENooXWgBmh

**Définition 21.77** ([603]).La **courbure totale** de la courbe  $\gamma: [a, b] \rightarrow \mathbb{R}^2$  est le nombre

$$K = \int_a^b \kappa(t) \|\gamma'(t)\| dt \quad (21.238)$$

EQooTIFWooQhOfd

où  $\kappa$  est la courbure signée définie en 21.73.Nous nommons aussi **nombre de tours** de  $\gamma$  le nombre  $\text{Turn}(\gamma) = K/2\pi$ .**Lemme 21.78.**

La courbure signée ne change pas sous reparamétrage positif, et change de signe sous reparamétrage négatif.

*Démonstration.* Soit la courbe  $\gamma: [a, b] \rightarrow \mathbb{R}^2$  et un difféomorphisme  $h: [a, b] \rightarrow [c, d]$ . Il s'agit d'intégrer la relation (21.227) en effectuant le changement de variables  $t = h(u)$  :

$$K_\gamma = \int_a^b \kappa_\gamma(t) \|\gamma'(t)\| dt \quad (21.239a)$$

$$= \int_c^d \underbrace{\kappa_\gamma(h(u))}_{\kappa_\sigma(u)} \|\gamma'(h(u))\| h'(u) du. \quad (21.239b)$$

En utilisant le fait que  $\sigma'(u) = \gamma'(h(u))h'(u)$  nous avons alors

$$K_\gamma = \int_c^d \kappa_\sigma(u) \left\| \frac{\sigma'(u)}{h'(u)} \right\| h'(u) du \quad (21.240a)$$

$$= \int_c^d \kappa_\sigma(u) \|\sigma'(u)\| \frac{h'(u)}{|h'(u)|} du \quad (21.240b)$$

$$= \text{sgn}(h') K_\sigma. \quad (21.240c)$$

□

LEMDEFooLPWJooAnWZjb

**Lemme-Définition 21.79** ([603]).Soit une courbe régulière  $\gamma: I \rightarrow \mathbb{R}^2$  de classe  $C^2$  et  $t_0$  dans l'intérieur de  $I$ . Soit  $\theta_0 \in \mathbb{R}$  tel que

$$\frac{\gamma'(t_0)}{\|\gamma'(t_0)\|} = (\cos(\theta_0), \sin(\theta_0)). \quad (21.241)$$

Alors il existe une unique fonction différentiable  $\theta: I \rightarrow \mathbb{R}$  telle que  $\theta(t_0) = \theta_0$  et

$$\frac{\gamma'(t)}{\|\gamma'(t)\|} = \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} \quad (21.242)$$

pour tout  $t \in I$ .Cette fonction est l'**angle** de  $\gamma$  déterminé par  $\theta_0$ .*Démonstration.* Soit  $\beta(t) = (t, 0)$ ; alors  $\beta'(t) = (1, 0)$  et nous avons

$$\gamma' \cdot \beta' = \gamma'_x \quad (21.243a)$$

$$\gamma' \cdot J\beta' = \gamma'_y. \quad (21.243b)$$

Par la proposition 21.71 il existe une unique fonction  $\theta: I \rightarrow \mathbb{R}$  telle que  $\theta(t_0) = \theta_0$  et

$$\begin{cases} \cos(\theta(t)) = \frac{\gamma' \cdot \beta'}{\|\gamma'\| \|\beta'\|} = \frac{\gamma'_x}{\|\gamma'\|} \\ \sin(\theta(t)) = \frac{\gamma' \cdot J\beta'}{\|\gamma'\| \|\beta'\|} = \frac{\gamma'_y}{\|\gamma'\|}. \end{cases} \quad (21.244a)$$

$$(21.244b)$$

Une telle fonction est bien celle que l'on demande ici.

□

LEMooWLAUooKetUiW

**Lemme 21.80.**

Si  $\gamma$  est une courbe régulière de classe  $C^2$ , alors sa courbure et son angle vérifient la relation

$$\theta'(t) = \|\gamma'(t)\| \kappa(t). \quad (21.245)$$

*Démonstration.* Par définition de l'angle (lemme 21.79) nous avons

$$\frac{\gamma'(t)}{\|\gamma'(t)\|} = (\cos \theta(t), \sin \theta(t)). \quad (21.246)$$

Dérivant cela,

$$\frac{\gamma''(t)}{\|\gamma'(t)\|} + \gamma'(t) \frac{d}{dt} \left( \frac{1}{\|\gamma'(t)\|} \right) = \theta'(t) \begin{pmatrix} -\sin(\theta(t)) \\ \cos(\theta(t)) \end{pmatrix} \quad (21.247a)$$

$$= \theta'(t) J \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} \quad (21.247b)$$

$$= \theta'(t) \frac{J\gamma'(t)}{\|\gamma'(t)\|}. \quad (21.247c)$$

Nous prenons le produit scalaire de cette égalité avec  $J\gamma'$  en tenant compte du fait que  $\gamma' \cdot J\gamma' = 0$  :

$$\frac{\gamma'' \cdot J\gamma'}{\|\gamma'\|} = \frac{\theta'}{\|\gamma'\|} J\gamma' \cdot J\gamma'. \quad (21.248)$$

En remarquant que  $Jv \cdot Jv = \|v\|^2$  nous trouvons  $\theta' \|\gamma'\| = \|\gamma'\|^2 \kappa_\gamma$  et donc

$$\theta'(t) = \|\gamma'(t)\| \kappa_\gamma(t), \quad (21.249)$$

ce qu'il fallait prouver.  $\square$

Ce lemme nous fournit la formule attendue pour la courbure totale.

**Lemme 21.81 ([603]).**

Soit une courbe régulière de classe  $C^2$   $\gamma: [a, b] \rightarrow \mathbb{R}^2$ . Sa courbure totale<sup>14</sup> est donnée par

$$K = \theta(b) - \theta(a). \quad (21.250)$$

*Démonstration.* Il suffit de remplacer dans la définition (21.238) de la courbure totale l'intégrande par son expression du lemme 21.80 :

$$K = \int_a^b \kappa(t) \|\gamma'(t)\| dt = \int_a^b \theta'(t) dt = \theta(b) - \theta(a) \quad (21.251)$$

par le théorème 14.288.  $\square$

**21.12.3 Degré, indice et homotopie**

La définition de l'indice d'un chemin par rapport à un point est 26.51.

DEFooTKBUooNVcheO

**Lemme-Définition 21.82.**

Soit une application continue  $\phi: S^1 \rightarrow S^1$ . Un **relèvement** de  $\phi$  est une application continue  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  une application continue telle que

$$\phi(\cos(t), \sin(t)) = \begin{pmatrix} \cos(\varphi(t)) \\ \sin(\varphi(t)) \end{pmatrix}. \quad (21.252)$$

Le **degré** est l'entier  $\deg(\phi)$  tel que

$$\varphi(2\pi) - \varphi(0) = 2 \deg(\phi) \pi. \quad (21.253)$$

Ce nombre ne dépend pas du choix du relèvement  $\varphi$ .

---

14. Définition 21.77.



*Démonstration.* Soient  $\varphi_1$  et  $\varphi_2$ , deux applications qui satisfont les contraintes. Nous avons une fonction continue  $n: S^1 \rightarrow \mathbb{R}$  telle que

$$\varphi_1(t) - \varphi_2(t) = 2\pi n(t). \quad (21.254)$$

La fonction  $n$  ne pouvant prendre que des valeurs entières et étant continue, elle est constante. Par conséquent

$$\varphi_1(2\pi) - \varphi_1(0) = \varphi_2(2\pi) - \varphi_2(0), \quad (21.255)$$

ce que nous voulions.  $\square$

DEFooOCUQooUAlbLo

### Définition 21.83.

Soit une courbe fermée  $\gamma: [0, L] \rightarrow \mathbb{R}^2$  de classe  $C^1$ . Nous posons

$$\Phi_\gamma(t) = \frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} \quad (21.256)$$

où  $\tilde{\gamma}(u) = \gamma(\frac{L}{2\pi}u)$ . L'**indice de rotation** de  $\gamma$  est le degré de  $\Phi_\gamma$ , c'est-à-dire

$$\text{Ind}(\gamma) = \deg(\Phi_\gamma). \quad (21.257)$$

Notons que dans cette définition,  $\Phi_\gamma$  n'est rien d'autre que le vecteur unitaire tangent à  $\gamma$ , ramené à  $[0, 2\pi]$ .

PROFoooXHSDoooDDnlJQ

### Proposition 21.84.

Pour une courbe fermée de classe <sup>15</sup>  $C^2$ , l'indice de rotation est égal au nombre de tours.

*Démonstration.* Soit  $\gamma: [0, L] \rightarrow \mathbb{R}^2$  une courbe vérifiant les hypothèses. Par définition,  $\text{Ind}(\gamma) = \deg(\Phi_\gamma)$  où

$$\Phi_\gamma(t) = \frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} = \begin{pmatrix} \cos \tilde{\theta}(t) \\ \sin \tilde{\theta}(t) \end{pmatrix} \quad (21.258)$$

EQoooGTAPoooDRCEPG

où nous avons noté  $\tilde{\theta}$  l'angle tournant <sup>16</sup> de  $\tilde{\gamma}$  et  $\theta$  celui de  $\gamma$ . Vu que  $\tilde{\theta}$  vérifie les hypothèses de la définition 21.82 nous pouvons calculer le degré de  $\Phi_\gamma$  par

$$\deg(\Phi_\gamma) = \frac{1}{2\pi} (\tilde{\theta}(2\pi) - \tilde{\theta}(0)) = \frac{1}{2\pi} \int_0^{2\pi} \tilde{\theta}'(s) ds. \quad (21.259)$$

EQoooKNCJoooTwTBMQ

Il faut trouver le lien entre  $\theta$  et  $\tilde{\theta}$ . Pour cela nous notons que

$$\frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} = \frac{\gamma'(\frac{L}{2\pi}t)}{\|\gamma'(\frac{L}{2\pi}t)\|}. \quad (21.260)$$

En comparant avec (21.258) il vient

$$\tilde{\theta}(t) = \theta\left(\frac{L}{2\pi}t\right) \quad (21.261)$$

et

$$\tilde{\theta}'(s) = \frac{L}{2\pi} \theta' \left( \frac{L}{2\pi} s \right). \quad (21.262)$$

Le changement de variables  $t = \frac{L}{2\pi}s$  est donc tout vu dans l'intégrale (21.259) :

$$\deg(\Phi_\gamma) = \frac{1}{2\pi} \int_0^{2\pi} \frac{L}{2\pi} \theta' \left( \frac{L}{2\pi} s \right) ds = \frac{1}{2\pi} \int_0^L \theta'(t) dt = \frac{1}{2\pi} \int_0^L \kappa_\gamma(t) \|\gamma'(t)\| dt = \text{Turn}(\gamma) \quad (21.263)$$

où nous avons aussi utilisé le lemme 21.80 qui donne le lien entre  $\theta'$  et  $\kappa$ .  $\square$

<sup>15</sup>. La dérivée seconde arrive dans la définition de la courbure ; il faudrait donc supposer au moins  $C^2$  pour avoir la continuité de la courbure.

<sup>16</sup>. Définition 21.79.

Les définitions à propos d'homotopie sont 20.77 et 20.78.

LEMooTDTAooCRSYzI

**Lemme 21.85** ([1]).

Soit un ouvert  $\Omega$  de  $\mathbb{C}$ . Soit un chemin continu  $\gamma: [a, b] \rightarrow \Omega$ . Soit une boule  $B$  contenant  $\gamma([a, b])$  ainsi que le segment joignant  $\gamma(a)$  et  $\gamma(b)$ .

Alors  $\gamma$  est homotope<sup>17</sup> au segment joignant  $\gamma(a)$  et  $\gamma(b)$ .

*Démonstration.* Notons  $d: [a, b] \rightarrow \Omega$  le segment de droite joignant  $\gamma(a)$  à  $\gamma(b)$ . Nous notons aussi  $l_{x,y}$  le segment joignant  $x$  à  $y$  et paramétré sur  $[0, 1]$  :

$$\begin{aligned} l_{x,y}: [0, 1] &\rightarrow \Omega \\ t &\mapsto (1-t)x + ty. \end{aligned} \quad (21.264)$$

L'homotopie entre  $\gamma$  et  $d$  est alors donnée par

$$\begin{aligned} H: [0, 1] \times [a, b] &\rightarrow \Omega \\ (s, t) &\mapsto l_{\gamma(t), d(t)}(s). \end{aligned} \quad (21.265)$$

□

LEMooQNFCooWorCxN

**Lemme 21.86** ([234]).

Soit un ouvert  $\Omega$  de  $\mathbb{C}$ . Tout chemin dans  $\Omega$  est homotope à extrémités fixées<sup>18</sup> dans  $\Omega$  à un chemin droit par morceaux.

*Démonstration.* Nous allons utiliser le lemme 9.23 en considérant  $S \subset [a, b]$  définie en disant que  $t \in S$  soit si  $t = a$  soit si le chemin  $\gamma: [a, t] \rightarrow \Omega$  est homotope à un chemin droit par morceaux à extrémités fixes.

Nous prouvons que  $S$  est inductif (définition 9.22).

- (1)  $a \in S$  Le point  $a$  est dans  $S$  par définition.
- (2) **Si**  $t \in S$  Supposons que  $t \in [a, b] \cap S$ , et cherchons  $s > t$  tel que  $[t, s] \subset S$ . Considérons une boule  $B = B(\gamma(t), \delta)$  incluse dans  $\Omega$ . Vu que  $\gamma$  est continue, l'application  $u \mapsto |\gamma(t) - \gamma(t+s)|$  est continue. Pour  $s = 0$ , cette fonction est nulle. Il existe  $\epsilon > 0$  tel que cette fonction reste plus petite que  $\delta$ , et donc tel que

$$\gamma([t, t+\epsilon]) \subset B(\gamma(t), \delta) \subset \Omega. \quad (21.266)$$

Le segment de droite  $l_{\gamma(t), \gamma(t+\epsilon)}$  est donc dans  $\Omega$  et la partie  $\gamma: [t, t+\epsilon] \rightarrow \Omega$  est homotope à extrémités fixée à ce segment (lemme 21.85).

En joignant ce segment aux segments de droites homotopes à  $\gamma([a, t])$ , nous avons des segments de droites homotopes à  $\gamma([a, s])$ . Donc  $[a, s] \subset S$ .

- (3) **Si**  $[a, t] \subset S$  Nous prouvons que  $t \in S$ . Nous considérons  $\delta > 0$  tel que  $B(\gamma(t), \delta) \subset \Omega$ . Nous considérons également  $\epsilon > 0$  tel que  $t - \epsilon > a$ . Par hypothèse,  $t - \epsilon \in S$ , et il existe une homotopie à extrémités fixées entre des segments de droites et  $\gamma: [a, t + \epsilon] \rightarrow \Omega$ . Nous complétons ces segments par le segment  $l_{\gamma(t-\epsilon), \gamma(t)}$  et voila.

Donc  $S$  est inductif, et nous en déduisons que  $S = [a, b]$ .

□

PROPooZIAKooHqtnZj

**Proposition 21.87** (Homotopie, degré et indice[603]).

Il y a deux résultats à ne pas confondre.

ITEMooLEHFooXEyTHY

- (1) Si  $\phi_i: S^1 \rightarrow S^1$  sont homotopes, alors  $\deg(\phi_1) = \deg(\phi_2)$ .
- (2) Si  $\gamma_i: [0, L] \rightarrow \mathbb{R}^2$  sont homotopes, alors  $\text{Ind}(\gamma_1) = \text{Ind}(\gamma_2)$ .

*Démonstration.* Nous allons décomposer la preuve en deux parties.

17. Définition 20.78.

18. Définition 20.78.

- (1) **Le degré pour les applications  $S^1 \rightarrow S^1$**  Soient deux applications homotopes  $\gamma_i: S^1 \rightarrow S^1$ . Si  $F$  est l'homotopie<sup>19</sup> entre  $\gamma_0$  et  $\gamma_1$ , nous posons  $\gamma_u(t) = F(u, t)$  qui est encore une courbe fermée  $\gamma_u: S^1 \rightarrow S^1$ . Nous pouvons donc considérer le degré de  $\gamma_u$ . C'est un entier  $n_u$  qui vérifie

$$\tilde{\gamma}_u(2\pi) - \tilde{\gamma}_u(0) = 2\pi n_u. \quad (21.267)$$

Le membre de gauche est une fonction continue de  $u$ ; dans le membre de droite  $n_u$  ne pouvant prendre que des valeurs entières, elle est alors constante.

- (2) **Indice de rotation pour des courbes dans  $\mathbb{R}^2$**  Soient maintenant  $\gamma_0$  et  $\gamma_1$  deux courbes homotopes de  $[0, L]$  dans  $\mathbb{R}^2$ . Par définition,  $\text{Ind}(\gamma_i) = \deg(\Phi_{\gamma_i})$ . Prouvons alors que  $\Phi_{\gamma_0}$  et homotope à  $\Phi_{\gamma_1}$ . De cette façon, la première partie de la preuve conclura à

$$\text{Ind}(\gamma_0) = \deg(\Phi_{\gamma_0}) = \deg(\Phi_{\gamma_1}) = \text{Ind}(\gamma_1). \quad (21.268)$$

Nous savons que

$$\Phi_\gamma = \frac{\gamma'(\frac{L}{2\pi}t)}{\|\gamma'(\frac{L}{2\pi}t)\|}, \quad (21.269)$$

donc en posant

$$\Phi_F(u, t) = \frac{\gamma'_u(\frac{L}{2\pi}t)}{\|\gamma'_u(\frac{L}{2\pi}t)\|} \quad (21.270)$$

nous avons une homotopie entre  $\Phi_{\gamma_0}$  et  $\Phi_{\gamma_1}$ .

□

THOooEQWOooBCRMMZ

### **Théorème 21.88 ([603]).**

Pour une courbe simple fermée de classe  $C^2$  dans  $\mathbb{R}^2$ ,

- (1) le nombre de tours<sup>20</sup> vaut  $\pm 1$ ,
- (2) la courbure totale vaut  $\pm 2\pi$ .

*Démonstration.* Soit une telle courbe  $\gamma$  et son image  $\Gamma$ . Notons tout de suite que les deux points sont équivalents parce que  $K = 2\pi \text{Turn}(\gamma)$ .

- (1) **Choix d'un point et d'une tangente**

Si  $\ell$  est une droite dans le plan, soit  $p$  le point de  $\Gamma$  le plus proche de  $\ell$ . Il n'est peut-être pas unique, mais la parallèle à  $\ell$  passant par  $p$  est une tangente à  $\Gamma$  telle que tout  $\Gamma$  se trouve d'un seul côté de  $\ell_p$ . Pour voir cela, il suffit de choisir un système d'axes pour lequel  $\ell$  est l'axe  $y = 0$ . La distance entre  $\ell$  et les points de  $\Gamma$  est donnée par  $\gamma_y$ , et donc les extrémums sont atteints là où  $\gamma'_y = 0$ , c'est-à-dire pour les points sur lesquels la tangente est parallèle à  $\ell$ .

Étant donné qu'il y a (au moins) un maximum et un minimum distincts, nous pouvons choisir le point  $p$  de telle sorte que pour tout point  $q \in \Gamma$ , l'angle du vecteur  $q - p$  soit entre  $\alpha_0$  et  $\alpha_0 + \pi$  et non entre  $\alpha_0 - \pi$  et  $\alpha_0$ . Ce choix revient à choisir  $p$  de telle sorte que  $\Gamma$  soit d'un côté ou de l'autre de  $\ell_p$ .

Pour simplifier les notations plus tard nous choisissons  $\ell_p$  horizontale, de telle sorte que  $\alpha_0 = 0$ , et que  $\Gamma$  est au dessus de  $\ell_p$ . Les angles des vecteurs  $q - p$  pour  $q \in \Gamma$  sont donc tous entre 0 et  $\pi$ .

- (2) **Définition de  $\Sigma$**

Soit  $L$  la longueur<sup>21</sup> de  $\gamma$  et  $\beta$ , un paramétrage de  $\gamma$  telle que  $\beta(0) = p$ . Nous considérons le triangle

$$\mathcal{T} = \{(t_1, t_2) \in \mathbb{R}^2 \text{ tel que } 0 \leq t_1 \leq t_2 \leq L\} \quad (21.271)$$

19. Définition 20.77.

20. Nombre de tours et courbure totale, définition 21.77.

21. Définition 21.3.

et l'application sécante  $\Sigma: \mathcal{T} \rightarrow S^1$  définie par

$$\Sigma(t_1, t_2) = \begin{cases} \frac{\beta'(t)}{\|\beta'(t)\|} & \text{si } t_1 = t_2 = t \\ -\frac{\beta'(0)}{\|\beta'(0)\|} & \text{si } t_1 = 0 \text{ et } t_2 = L \\ \frac{\beta(t_2) - \beta(t_1)}{\|\beta(t_2) - \beta(t_1)\|} & \text{sinon.} \end{cases} \quad (21.272)$$

(3) **Continuité de  $\Sigma$**  Nous devons prouver les limites<sup>22</sup> suivantes :

$$\lim_{\substack{(t_1, t_2) \rightarrow (t, t) \\ 0 \leq t_1 \leq t \leq t_2 \leq L}} \frac{\beta(t_1) - \beta(t_2)}{\|\beta(t_1) - \beta(t_2)\|} = \frac{\beta'(t)}{\|\beta'(t)\|} \quad \text{SUBEQooDXJDooQLscBD} \quad (21.273a)$$

$$\lim_{\substack{(t_1, t_2) \rightarrow (0, L) \\ 0 \leq t_1 \leq t_2 \leq L}} \frac{\beta(t_1) - \beta(t_2)}{\|\beta(t_1) - \beta(t_2)\|} = -\frac{\beta'(0)}{\|\beta'(0)\|} \quad \text{SUBEQooOXGSooXHEHHh} \quad (21.273b)$$

Nous commençons par (21.273a) en multipliant et divisant par  $t_2 - t_1$  :

$$\frac{\beta(t_2) - \beta(t_1)}{\|\beta(t_2) - \beta(t_1)\|} = \frac{\beta(t_2) - \beta(t_1)}{t_2 - t_1} \frac{t_2 - t_1}{\|\beta(t_2) - \beta(t_1)\|}. \quad (21.274)$$

Si chacune des limites des deux facteurs existent dans  $\mathbb{R}$ , la limite du produit sera le produit des limites. Pour le premier facteur nous développons  $\beta(t_2)$  autour de  $t = t_1$  via la formulation (12.1408) :

$$\beta(t_2) = \beta(t_1) + (t_2 - t_1)\beta'(t_1) + \alpha(t_2 - t_1) \quad (21.275)$$

où  $\alpha$  est une fonction vérifiant  $\lim_{t \rightarrow 0} \alpha(t)/t = 0$ . Nous avons à calculer la limite de

$$\frac{(t_2 - t_1)\beta'(t_1) + \alpha(t_2 - t_1)}{t_2 - t_1} = \beta'(t_1) + \frac{\alpha(t_2 - t_1)}{t_2 - t_1}. \quad (21.276)$$

Prendre la limite  $(t_1, t_2) \rightarrow (t, t)$  donne bien  $\beta'(t)$  parce que  $\beta$  est de classe  $C^1$ . En ce qui concerne la limite de la deuxième partie, nous allons la faire plus en détail pour la limite (21.273b).

La limite (21.273b). Nous prenons la prolongation périodique de  $\beta$ . Alors si  $t_2 = L - \epsilon$  nous pouvons écrire  $\beta(-\epsilon)$  au lieu de  $\beta(t_2)$ . Nous développons  $\beta(t_1)$  autour de  $t = -\epsilon$  (parce que  $t_1$  est petit) :

$$\beta(t_1) = \beta(-\epsilon) + (t_1 + \epsilon)\beta'(-\epsilon) + \alpha(t_1 + \epsilon). \quad (21.277)$$

Après multiplication et division par  $t_1 + \epsilon$ , la première limite à calculer est celle de

$$\frac{\beta(-\epsilon) - \beta(t_1)}{t_1 + \epsilon} = -\beta'(-\epsilon) + \frac{\alpha(t_1 + \epsilon)}{t_1 + \epsilon}, \quad (21.278)$$

pour  $(\epsilon, t_1) \rightarrow (0, 0)$ . Cela donne bien  $-\beta'(0)$ . La seconde limite à calculer est celle de

$$\frac{t_1 + \epsilon}{\|\beta(-\epsilon) - \beta(t_1)\|} = \frac{t_1 + \epsilon}{\|-(t_1 + \epsilon)\beta'(-\epsilon) - \alpha(t_1 + \epsilon)\|} = \frac{t_1 + \epsilon}{\|(t_1 + \epsilon)\beta'(-\epsilon) + \alpha(t_1 + \epsilon)\|}. \quad (21.279)$$

Nous calculons la limite de l'inverse (qui, si elle est non nulle donnera la réponse en inversant à nouveau) en nous souvenant de la formule

$$|a - b| \leq |a + b| \leq |a| + |b|. \quad (21.280)$$

Nous avons l'encadrement

$$\left| \frac{(t_1 + \epsilon)\beta'(-\epsilon) - \alpha(t_1 + \epsilon)}{t_1 + \epsilon} \right| \leq \frac{\|(t_1 + \epsilon)\beta'(-\epsilon) + \alpha(t_1 + \epsilon)\|}{t_1 + \epsilon} \quad (21.281a)$$

$$\leq \frac{|(t_1 + \epsilon)\beta'(-\epsilon)| + |\alpha(t_1 + \epsilon)|}{t_1 + \epsilon} \quad (21.281b)$$

Les limites des deux extrêmes existent et valent  $\beta'(0)$  ; la règle de l'étau 12.246 conclu.

22. Limite au sens de la définition 7.135, en sachant qu'elle est unique par la proposition 7.138.

- (4) **Deux chemins homotopes** Nous considérons dans  $\mathcal{T}$  les points  $A = (0, 0)$ ,  $B = (0, L)$  et  $C = (L, L)$ . Nous allons considérer le chemin direct de  $A$  à  $C$  et celui passant via  $B$ . Et comme ces chemins doivent être paramétrés de 0 à  $2\pi$ , il faut faire un peu attention. Nous définissons les chemins

$$\sigma_i: S^1 \rightarrow S^1 \quad (21.282)$$

par

$$\sigma_1(t) = \Sigma\left(\frac{L}{2\pi}t, \frac{L}{2\pi}t\right) = \frac{\beta'(tL/2\pi)}{\|\beta'(tL/2\pi)\|} = \frac{\tilde{\beta}'(t)}{\|\tilde{\beta}'(t)\|} = \Phi_\beta \quad (21.283a)$$

$$\sigma_2(t) = \begin{cases} \Sigma(0, \frac{L}{\pi}t) & \text{si } 0 \leq t \leq \pi \\ \Sigma(\frac{L}{\pi}(t - \pi), L) & \text{si } \pi \leq t \leq 2\pi \end{cases} \quad (21.283b)$$

où nous avons repris les notations de la définition 21.83. Notons que pour  $\sigma_2$ , en  $t = \pi$  les deux expressions donnent

$$\sigma_2(\pi) = \Sigma(0, L) = -\frac{\beta'(0)}{\|\beta'(0)\|}. \quad (21.284)$$

Ce sont des chemins fermés parce que

$$\sigma_1(2\pi) = \Sigma(L, L) = \frac{\beta'(L)}{\|\beta'(L)\|} = \frac{\beta'(0)}{\|\beta'(0)\|} = \Sigma(0, 0) = \sigma_1(0). \quad (21.285)$$

Notons que dans toutes ces définitions et calculs, nous avons utilisé de façon assez cruciale la définition 21.70 pour définir la dérivée de  $\beta$  en  $t = 0$ .

Les chemins  $\sigma_1$  et  $\sigma_2$  sont homotopes par construction.

- (5) **Indices et degrés** La proposition 21.87(1) nous donne  $\deg(\sigma_1) = \deg(\sigma_2)$ . Nous avons alors la chaîne d'égalités

$$\deg(\sigma_2) = \deg(\sigma_1) = \deg(\Phi_\beta) = \text{Ind}(\beta) = \text{Ind}(\gamma) = \text{Turn}(\gamma) \quad (21.286)$$

où les justifications sont :

- (5a)  $\deg(\sigma_2) = \deg(\sigma_1)$  par homotopie : proposition 21.87.
- (5b)  $\deg(\sigma_1) = \deg(\Phi_\beta)$  parce que  $\sigma_1 = \Phi_\beta$ .
- (5c)  $\deg(\Phi_\beta) = \text{Ind}(\beta)$  par définition 21.83 de l'indice.
- (5d)  $\text{Ind}(\beta) = \text{Ind}(\gamma)$  par invariance de l'indice sous reparamétrage.
- (5e)  $\text{Ind}(\gamma) = \text{Turn}(\gamma)$  par la proposition 21.84.

Il nous reste à montrer que  $\deg(\sigma_2) = \pm 1$ .

- (6) **La géométrie de  $\sigma_2$**  Notons que par définition les valeurs de  $\sigma_2(t)$  sont les vecteurs (unitaires) joignant  $p$  aux points de  $\Gamma$  lorsque  $0 \leq t \leq \pi$ , et les vecteurs inverses pour  $\pi \leq t \leq 2\pi$ . Plus précisément nous avons, si  $0 < a < \pi$  :

$$\sigma_2(a) = \Sigma(0, \frac{L}{\pi}a) = \frac{\beta(\frac{L}{\pi}a) - \beta(0)}{\|\beta(\frac{L}{\pi}a) - \beta(0)\|} \quad (21.287)$$

et

$$\sigma_2(\pi + a) = \Sigma(\frac{L}{\pi}a, L) = \frac{\beta(L) - \beta(\frac{L}{\pi}a)}{\|\beta(L) - \beta(\frac{L}{\pi}a)\|}. \quad (21.288)$$

En sachant que  $\beta(0) = \beta(L)$  nous avons alors

$$\sigma_2(\pi + a) = -\sigma_2(a). \quad (21.289)$$

Notons aussi que le fait que  $\gamma$  soit une courbe simple assure que le numérateur et le dénominateur de  $\sigma_2$  ne s'annulent pas autrement que pour  $t = L$  ou  $t = 0$ .

(7) Un relèvement pour  $\sigma_2$  Ces propriétés motivent cette idée pour le relèvement de  $\sigma_2$  :

$$\begin{aligned} \theta: [0, 2\pi[ &\rightarrow [0, 2\pi[ \\ t &\mapsto \text{l'angle du vecteur } \sigma_2(t), \end{aligned} \quad (21.290)$$

avec  $\theta(2\pi)$  défini par continuité. Nous allons cependant voir, en étant plus prudent, que cette définition n'assure pas la continuité (surtout en  $t = \pi$ ).

Soyons donc plus prudent et construisons  $\theta$  petit à petit.

Étant donné que  $\ell_p$  est tangente à  $\Gamma$  au point  $p$ , la droite  $\ell_p$  est parallèle à  $\beta'(0)$ , et l'angle entre  $\ell_p$  et  $\beta'(0)$  est soit 0 soit  $\pi$ . Donc  $\theta(0)$  devrait valoir soit 0 soit  $\pi$ .

Nous commençons par définir ceci :

$$\begin{aligned} \theta: [0, \pi[ &\rightarrow [0, \pi] \\ t &\mapsto \text{angle de } \sigma_2(t) = \arccos(\sigma_2(t)_x). \end{aligned} \quad (21.291)$$

C'est le choix d'avoir  $\ell_p$  horizontale et  $\Gamma$  au dessus de  $\ell_p$  qui nous assure que pour tout  $t \in [0, \pi]$ , l'angle de  $\sigma_2(t)$  peut être choisi entre 0 et  $\pi$ . De plus cette fonction est continue en tant que partie de la fonction  $\arccos: [-1, 1] \rightarrow [0, \pi]$  qui est elle-même continue par le lemme 7.277.

Le fait que  $\theta$  soit continue est assuré par le fait que  $\sigma_2$  est  $C^\infty$ .

(7a) Si  $\theta(0) = 0$

Cela correspond à la situation des vecteurs rouges sur la figure 21.11.

Vu que  $\sigma_2(\pi) = -\sigma_2(0)$ , l'angle de  $\sigma_2$  en  $t = \pi$  est le supplémentaire de celui en  $t = 0$ . Mais pour  $0 \leq t < \pi$ ,  $\theta(t)$  prend ses valeurs entre 0 et  $\pi$ , le seul supplémentaire de 0 à être disponible est  $\theta(\pi) = \pi$  (et non  $\theta(\pi) = -\pi$  par exemple). Nous définissons donc  $\theta(\pi) = \pi$  pour la continuité.

En ce qui concerne les  $\pi < t < 2\pi$  nous savons que  $\sigma_2(\pi + a) = -\sigma_2(a)$ . Les angles sont donc les supplémentaires de ceux pour  $0 \leq t \leq \pi$ . Pour assurer la continuité en  $t = \pi$  nous sélectionnons la place  $]\pi, 2\pi]$  ; et nous définissons

$$\begin{aligned} \theta: ]\pi, 2\pi[ &\rightarrow ]\pi, 2\pi] \\ t &\mapsto \text{angle de } \sigma_2(t). \end{aligned} \quad (21.292)$$

Le nombre  $\theta(2\pi)$  est défini par continuité. Il doit valoir  $\theta(2\pi) = 2\pi$ .

La fonction  $\theta$  ainsi définie est un relèvement pour  $\sigma_2$ , et le degré peut être calculé :

$$\deg(\sigma_2) = \frac{1}{2\pi}(\theta(2\pi) - \theta(0)) = 1. \quad (21.293)$$

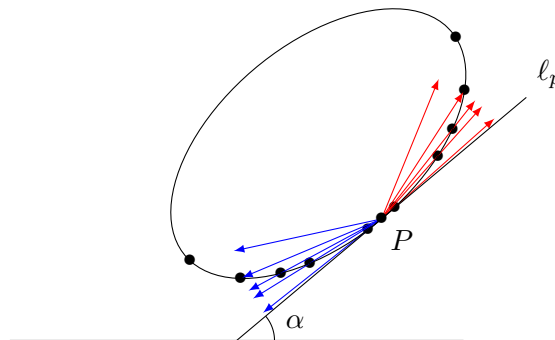


FIGURE 21.11: Les vecteurs représentant  $\sigma_2$  dans le cas où  $\beta'(0)$  est dans le sens de  $\ell_p$  ou dans le sens inverse. Pour le sport nous avons dessiné la situation avec une droite  $\ell$  quelconque plutôt que horizontale.

(7b) **Si**  $\theta(0) = \pi$  Cela correspond à la situation des vecteurs bleus sur la figure 21.11.

Alors  $\theta(\pi) = 0$  est obligatoire parce qu'il doit être supplémentaire à  $\theta(0)$ . Les angles atteints par  $\sigma_2(t)$  pour  $t \in ]\pi, 2\pi[$  sont encore les complémentaires, mais cette fois la continuité en  $t = \pi$  nous impose de les chercher dans  $]-\pi, 0]$  et nous définissons

$$\begin{aligned} \theta: ]\pi, 2\pi[ &\rightarrow ]-\pi, 0] \\ t &\mapsto \text{angle de } \sigma_2(t). \end{aligned} \quad (21.294)$$

Par continuité nous devons avoir  $\theta(2\pi) = \theta(\pi) - \pi$  et donc  $\theta(2\pi) = -\pi$ .

Le degré de  $\sigma_2$  est alors

$$\deg(\sigma_2) = \frac{1}{2\pi}(\theta(2\pi) - \theta(0)) = -1. \quad (21.295)$$

#### (8) **Conclusion**

Nous avons montré que le degré de  $\sigma_2$  est 1 ou  $-1$ , et en remontant les égalités (21.286) nous déduisons que  $\text{Turn}(\gamma) = \pm 1$ .

□

## 21.13 Courbes fermées planes

### 21.13.1 Cercle circonscrit

La proposition suivante est dans le même esprit que l'ellipse de John-Loewer<sup>23</sup>.

PROPDÉFooCWESooVbDven

**Proposition-Définition 21.89** (Cercle circonscrit[670, 1]).

Soit une courbe fermée simple et continue  $\gamma: [0, 1] \rightarrow \mathbb{R}^2$ . Soit  $\Gamma$  son image. Il existe un unique cercle de rayon minimum contenant  $\Gamma$ . Ce cercle est le **cercle circonscrit** à  $\gamma$ .

Nous nommons  $C$  ce cercle circonscrit à  $\gamma$ . Il a les propriétés suivantes :

- (1) L'intersection  $C \cap \Gamma$  contient au moins deux points distincts.
- (2) Tout arc de  $C$  plus long que sa moitié intersecte  $\Gamma$ .

*Démonstration.* Division de la preuve.

- (1) **Existence** L'application  $\gamma$  étant continue, l'ensemble  $\Gamma$  est compact (théorème 7.274). Nous considérons l'ensemble  $Q$  des formes quadratiques de la forme

$$q_{a,r}(x) = \|a - x\|^2 - r^2 \quad (21.296)$$

avec  $a \in \mathbb{R}^2$  et  $r \in \mathbb{R}$ . Nous mettons sur cet ensemble la topologie de  $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$ . Le nombre  $q_{a,r}(x)$  est continu en  $a$ ,  $r$  et  $x$ . Soit  $A$  l'ensemble des formes quadratiques de cette forme et vérifiant

$$q(\gamma(t)) \leq 0 \quad (21.297)$$

pour tout  $t \in [0, 1]$ .

Cet ensemble  $A$  est non vide parce que  $\Gamma$  est compact et donc borné ; il existe donc une boule qui contient  $\Gamma$  en son intérieur.

Montrons que  $A$  est fermé dans  $Q$ . Si  $q_{a,r} \notin A$  alors il existe  $t_0$  tel que  $q_{a,r}(\gamma(t_0)) > 0$ . Par continuité, il existe un voisinage de  $(a, r)$  dans  $\mathbb{R}^3$  et donc de  $q_{a,r}$  dans  $Q$  tel que  $q_{a',r'}(\gamma(t_0))$  reste strictement positif pour tout  $(a', r')$  dans ce voisinage. Le complémentaire de  $A$  dans  $Q$  est ouvert. Donc  $A$  est fermé.

L'ensemble

$$B = \{r \geq 0 \text{ tel que } \exists a \text{ tel que } q_{a,r} \in A\} \quad (21.298)$$

est fermé<sup>24</sup>. La partie  $B$  est évidemment borné vers le bas. De plus  $r = 0$  n'est pas dans cet ensemble. Il possède donc un minimum strictement positif. Le cercle correspondant donne l'existence.

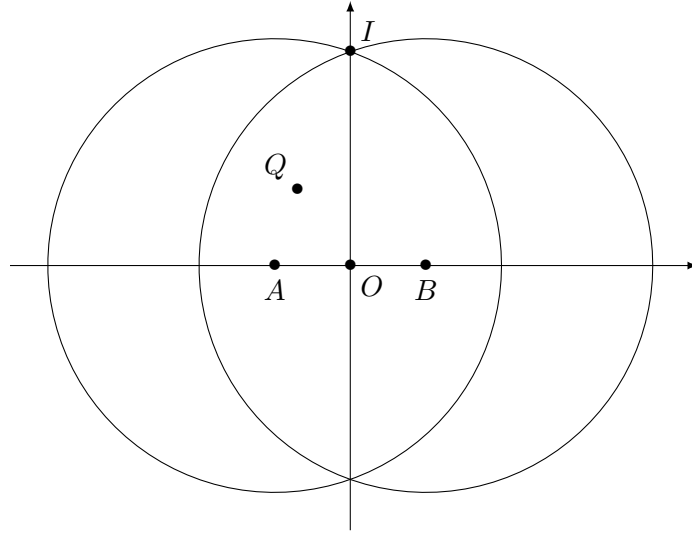
23. Proposition 17.122.

24. Lemme 7.375.

(2) **Unicité**

En ce qui concerne l'unicité, si  $\Gamma$  est contenu dans les boules  $B(a, R)$  et  $B(b, R)$  alors

$$\Gamma \subset B(a, R) \cap B(b, R). \quad (21.299)$$



Nous choisissons les axes comme indiqué sur le dessin et nous montrons que l'intersection est dans le cercle de centre  $O$  et de rayon  $\|OI\| < R$ . Soit  $Q$  un point de l'intersection ; par symétrie il est suffisant de supposer  $Q_x < 0$  et  $Q_y > 0$ . Vu que  $Q$  est dans le cercle de centre  $B$  et de rayon  $R$ , il doit satisfaire

$$(B_x - Q_x)^2 + Q_y^2 < R. \quad (21.300)$$

D'autre part le point  $I$  est d'abscisse  $I_x = 0$  et d'ordonnée donnée par  $I_y^2 = R^2 - B_x^2$ .

Nous devons prouver que  $Q_x^2 + Q_y^2 \leq I_y^2$ . Il s'agit simplement de calculer

$$Q_x^2 + Q_y^2 \leq Q_x^2 + R^2 - B_x^2 + 2B_x Q_x - Q_x^2 = R^2 - B_x^2 + 2B_x Q_x \leq R^2 - B_x^2 = I_y^2 \quad (21.301)$$

parce que  $B_x > 0$  et  $Q_x < 0$ .

Nous concluons que  $\Gamma$  est inclus dans un cercle de rayon plus petit que  $R$  et donc que  $R$  n'est pas minimum. D'où l'unicité.

(3) **Au moins deux intersections**

Nous nommons  $C$  le cercle circonscrit à  $\gamma$ , et nous écrivons, pour  $p \in \Gamma$

$$r(p) = d(p, C) \quad (21.302)$$

la distance entre  $p$  et  $C$ . Cela est une fonction continue sur le compact  $\Gamma$ . Elle atteint donc ses bornes sur  $\Gamma$ .

(3a) **Impossible d'avoir zéro intersection**

Si  $C \cap \Gamma = \emptyset$  alors  $r(p) > 0$  pour tout  $p$  et le minimum est également strictement positif. Soit  $r_0$  ce minimum :

$$r_0 = \min\{d(p, C) \text{ tel que } p \in \Gamma\}. \quad (21.303)$$

Alors le cercle  $C_2$  même centre que  $C$  mais de rayon  $r_0/2$  n'intersecte pas non plus  $\Gamma$ . En effet soient  $p \in \Gamma$  et  $s \in C_2$ . Nous notons  $s'$  le point de  $C$  situé sur la demi-droite  $ps$ . Nous avons

$$d(p, s) \geq d(p, s') - \frac{r_0}{2} \geq d(p, C) - \frac{r_0}{2} \quad (21.304a)$$

$$\geq r_0 - \frac{r_0}{2} = \frac{r_0}{2} \quad (21.304b)$$

$$> 0. \quad (21.304c)$$

Cela étant valable pour tout  $s \in C_2$  (et  $d$  étant continue en  $s$  sur le compact  $C_2$ ), nous en déduisons que  $d(p, C_2) > 0$  pour tout  $p \in \Gamma$ . Donc  $\Gamma \cap C_2 = \emptyset$ .



(3b) **Impossible d'avoir une seule intersection**

Si  $C \cap \Gamma = \{q\}$  alors un choix d'axe place le centre de  $C$  en  $(0, 0)$ , le point  $q$  en  $(1, 0)$  et fixe le rayon de  $C$  à 1. Vu que  $\Gamma$  ne coupe  $C$  qu'un seul point, situé tout à droite, il devrait être possible de décaler un peu  $C$  pour le « décoller » de  $q$  tout en continuant à ne pas toucher  $\Gamma$  en les autres points.

C'est ce que nous allons faire maintenant.

Pour  $p \in \Gamma$  nous notons  $r(p)$  la distance entre  $p$  et le point de  $C$  situé sur sa gauche, c'est-à-dire, si  $p = (p_x, p_y)$ ,

$$r(p) = \sqrt{1 - p_y^2} + p_x. \quad (21.305) \quad \text{EQooCWRDooRzwgKi}$$

Cela est encore une fonction continue sur  $\Gamma$  qui atteint son minimum valant  $r_0$ . Nous considérons maintenant le cercle  $C_2$  de centre  $(\frac{r_0}{2}, 1)$ , et nous prouvons que  $\Gamma$  est dans  $C_2$  (au sens dans la boule dont la frontière est  $C_2$ ) et que  $C_2 \cap \Gamma = \emptyset$ . Donc  $C_2$  retombe dans un cas que nous avons prouvé être impossible.

Soit  $p \in \Gamma$ . Si  $p \in C_2$ , alors  $p - (r_0/2, 0) \in C$ . Cela est impossible parce que la distance entre  $p$  et le point à sa gauche sur  $C$  est au moins  $r_0$ . Donc  $C_2 \cap \Gamma = \emptyset$ .

Le point  $q$  est dans  $C_2$  parce que  $(1, 0) \in B((r_0/2, 0), 1)$ . Par continuité de  $C_2$  et de  $\Gamma$ , tous les points de  $\Gamma$  sont dans  $C_2$ . Sinon il y aurait une intersection.

(3c) **Si on en veut plus** Nous avons fini de démontrer qu'il y a au moins deux intersections.

Pour le sport, notez que tout nombre de points d'intersection entre  $C$  et  $\Gamma$  est possible à partir de 2. Pour en avoir deux, prendre une ellipse, et pour en avoir plus, prendre des polynômes dont les angles sont un peu modifiés de façon à rester  $C^1$ .

(4) **Intersection avec les demi-arcs**

Supposons, en fixant encore les axes, que le cercle circonscrit soit encore centré en  $(0, 0)$  et que  $\Gamma$  n'intersecte pas le demi-cercle  $x < 0$ . Alors pour tout  $p \in \Gamma$  la distance entre  $p$  et ce demi-cercle est strictement positive. Il y a un minimum  $r_0$ . En décalant le centre du cercle de  $r_0/2$  vers la droite, nous obtenons un nouveau cercle contenant  $\Gamma$  mais ne l'intersectant pas.

□

## 21.13.2 Description locale

DEFooVQODooJSNYLw

**Définition 21.90.**

Une courbe plane différentiable est **convexe** si son graphe est en tout point d'un seul côté de sa tangente.

LEMooGEVEooHxPTMO

**Lemme 21.91** ([1]).

Soit une courbe simple, convexe  $\gamma: [0, 1] \rightarrow \mathbb{R}^2$  que nous supposons être de classe  $C^1$ . Nous notons  $\Gamma$  l'image de  $\gamma$ . Alors  $\Gamma$  est localement le graphe d'une fonction convexe (définition 17.76).

*Démonstration.* Soit  $p \in \Gamma$  et  $l_p$  la tangente à  $\Gamma$  en  $p$ . Nous considérons un système d'axe centré en  $p$  de telle sorte que  $l_p$  soit l'axe des abscisses et l'axe des ordonnées soit dirigé de tel manière que  $\Gamma$  se trouve dans la partie  $y > 0$ . De plus nous paramétrons  $\gamma$  de telle sorte à avoir  $\gamma(0) = p = (0, 0)$ .

Vu que  $l_p \equiv y = 0$  est la tangente à  $\Gamma$  nous avons  $\gamma'_y(0) = 0$  et  $\gamma'_x(0) > 0$ . Nous en déduisons que  $\gamma'_x(t) > 0$  pour tout  $t \in B(0, \delta)$  pour  $\delta$  suffisamment petit. Nous posons alors

$$g(x) = \gamma_y(\gamma_x^{-1}(x)) \quad (21.306)$$

qui est bien définie parce que  $\gamma_x$  est une bijection entre  $B(0, \delta)$  et son image. La fonction  $g$  est continue par le théorème de la bijection 12.53, et même dérivable par la proposition 12.211. De plus, vu la formule (12.528), la fonction  $g^{-1}$  est de classe  $C^1$  parce que  $(g^{-1})'$  est une composée d'applications continues.

- (1) **Si  $g$  est  $C^2$**  Dans ce cas,  $g''$  ne peut pas changer de signe, sinon la tangente coupe le graphe. Par positivité de  $g$  (et le fait que  $g(0) = g'(0) = 0$ ), il n'est pas possible d'avoir  $g'' < 0$  partout. Donc  $g'' \geq 0$  partout. Cela prouve que  $g$  est convexe par la caractérisation 17.84.
- (2) **Si  $g$  est seulement de classe  $C^1$**  Le graphe de  $g$  correspond au graphe de  $\Gamma$ . Nous montrons que  $g$  est convexe en utilisant la caractérisation de la proposition 17.89. La tangente au graphe de  $g$  en  $x = x_0$ , que nous notons  $l_0$ , est la tangente à  $\Gamma$  en  $t = \gamma_x^{-1}(x_0)$ . Le graphe de  $g$ , qui est une partie de  $\Gamma$  se trouve donc d'un seul côté de  $l_0$ . Nous nous restreignons  $g$  à un compact  $I$  et nous considérons la fonction

$$d_a(x) = g(x) - l_a(x) \quad (21.307)$$

qui donne la distance entre le graphe de  $g$  et la tangente à  $g$  en  $x = a$ . Cela est une fonction continue en  $x$  et en  $a$ . Le graphe de  $g$  est au dessus de la tangente en  $x = 0$  (par construction des axes). Supposons que le graphe de  $g$  soit en dessous de la tangente en  $x = x_2$ . Alors nous avons, pour tout  $x$  :

$$\begin{cases} d_0(x) \geq 0 \\ d_{x_1}(x) \leq 0. \end{cases} \quad (21.308a)$$

$$(21.308b)$$

Nous posons

$$s(a) = \sup_{x \in I} d_a(x), \quad (21.309)$$

qui est une fonction continue par la proposition 12.35. L'application  $s$  vérifie  $s(0) \geq 0$  et  $s(x_2) \leq 0$ . Il existe donc  $m \in [0, x_2]$  tel que  $s(m) = 0$ . À ce moment nous avons  $g(x) = l_m(x)$  pour tout  $x \in I$  et donc  $g$  est une droite, et en réalité toutes les inégalités sont des égalités. La fonction  $g$  est alors bien convexe (mais pas strictement). □

### 21.13.3 Enveloppe convexe

PROPOoWZITooTFiWsi

**Proposition 21.92** ([671, 1]).

Soit une courbe simple, fermée et convexe  $\gamma: [0, 1] \rightarrow \mathbb{R}^2$  que nous supposons être de classe  $C^2$ . Nous notons  $\Gamma$  l'image de  $\gamma$ . Alors il existe un convexe  $D$  tel que  $\partial D = \Gamma$ .

*Démonstration.* Pour  $p \in \Gamma$  nous notons  $l_p$  la tangente à  $\Gamma$  en  $p$  et  $H_p$  le demi-plan (fermé) contenant  $\Gamma$ . Nous posons

$$D = \bigcap_{p \in \Gamma} H_p. \quad (21.310) \quad \text{EQooDYFTooCHFBsD}$$

Cet ensemble est convexe comme intersection de convexes<sup>25</sup> et fermé comme intersection de fermés<sup>26</sup>. Nous prouvons que  $\Gamma = \partial D$ .

L'inclusion  $\Gamma \subset \partial D$  est la plus facile. Si  $p \in \Gamma$  alors  $p$  est dans chacun des  $H_q$  et donc dans  $D$ . De plus tout voisinage de  $p$  contient des points en dehors de  $H_p$ , donc  $p$  n'est pas dans l'intérieur de  $D$ . Ce dernier étant fermé, un point hors de l'intérieur est sur le bord. Ergo  $p \in \partial D$ .

Pour l'inclusion inverse, soit  $p \in \partial D$ .

- (1) **Il existe  $q$  tel que  $p \in l_q$**  Vu que  $D$  est fermé, le point  $p$  est dans  $D$ , et donc dans tous les  $H_q$ . Supposons qu'il soit dans l'intérieur de tous les  $H_q$ . Alors nous considérons la fonction

$$r(q) = \frac{d(p, l_q)}{2} \quad (21.311)$$

définie sur  $\Gamma$ . C'est une fonction continue<sup>27</sup> strictement positive définie sur le compact  $\Gamma$  qui possède donc un minimum strictement positif. Si  $r_0$  est ce minimum, alors  $B(p, r_0)$  est incluse à tous les  $H_q$ , ce qui ferait que  $p$  est à l'intérieur de  $D$ . Nous concluons à l'existence de  $q \in \Gamma$  tel que  $p \in l_q$ .

25. Proposition 7.209.

26. Lemme 7.6(1).

27. L'équation de la droite  $l_q$  a des coefficients continus parce que  $\gamma$  est de classe  $C^1$ .

- (2) **Le point où ça décolle** Nous supposons que  $p \notin \Gamma$ , sinon ce serait trop facile. Nous paramétrons  $\gamma$  de telle sorte à avoir  $q = \gamma(0)$  et nous posons

$$T = \{t \in \mathbb{R}^+ \text{ tel que } l_{\gamma(t)} = l_q\}. \quad (21.312)$$

Cela est un fermé dans  $\mathbb{R}$  parce que  $\gamma$  est de classe  $C^1$ . Nous posons  $t_0 = \inf(T^c)$  et pour tout  $\epsilon$  suffisamment petit,

$$\begin{cases} l_{\gamma(t_0+\epsilon)} \neq l_q, \\ l_{\gamma(t_0-\epsilon)} = l_q. \end{cases} \quad (21.313a)$$

$$(21.313b)$$

La seconde est parce que si  $l_{\gamma(t_0-\epsilon)} \neq l_q$  nous aurions  $t_0 - \epsilon \in T^c$ . Soit  $r = \gamma(t_0)$ ; nous avons  $l_r = l_q$  parce que si  $l_r \neq l_q$  alors par continuité de  $\gamma'$  nous aurions  $l_{r'} \neq l_q$  pour tout  $r' \in \gamma(B(t_0, \delta))$ .

- (3) **Graphes d'une fonction strictement convexe** En suivant le lemme 21.91, l'ensemble  $\Gamma$  est localement (autour de  $r$ ) le graphe d'une fonction convexe au-dessus de  $l_r$ . Soit  $g: B(0, \delta) \rightarrow \mathbb{R}^2$  cette fonction convexe.

Soit  $\epsilon > 0$ . Si  $g''(x) = 0$  sur  $]0, \epsilon]$  alors  $g'$  y est constante. Mais  $g'(0) = 0$ , ce qui signifierait que sur  $[0, \epsilon]$  nous ayons  $g'(x) = 0$  et donc  $g(x) = 0$ . Cela ferait que  $l_{r'} \equiv y = 0$  pour tout  $r'$  de la forme  $(x, 0)$  avec  $x \in [0, \epsilon]$  (qui sont des points de  $\Gamma$ ). Cela est en contradiction avec la définition de  $r$ . Donc il existe un point  $x \in [0, \epsilon[$  tel que  $g''(x) > 0$ .

Rappelons que  $p \in l_q$ , ce qui fait que  $p$  a pour coordonnées  $(p_x, 0)$ . Nous restreignons  $\delta$  et  $\epsilon$  de telle sorte que  $p_x$  soit plus grand à la fois que  $\epsilon$  et  $\delta$ .

Il existe donc un intervalle  $[a, b]$  avec  $a, b \geq 0$  et  $a, b < p_x$  sur lequel  $g$  est strictement convexe.

- (4) **La tangente qui tue**

En particulier  $g(b) > 0$ , de telle sorte que le théorème des accroissements finis 12.216(1) nous donne l'existence de  $m \in [0, b]$  tel que la tangente à  $g$  en  $x = m$  est parallèle au segment joignant  $(0, 0)$  à  $(b, g(b))$ . Cette tangente, que nous nommons  $l_m$ , est en dessous de la corde, par stricte convexité. En particulier, son point d'intersection avec  $y = 0$  est strictement entre 0 et  $m$ .

L'ensemble  $D$  est d'un seul côté de  $l_m$ . Ce côté est forcément celui de  $q = (0, 0)$  (parce que  $q \in \Gamma \subset D$ ), et donc les points de coordonnées  $(x, 0)$  avec  $x > m$  ne sont pas dans  $D$ . Pas de chance,  $p$  est un point de ce type.

- (5) **La contradiction** Nous avons prouvé que si  $p \in \partial D \setminus \Gamma$  alors  $p$  n'est pas dans  $D$ , ce qui est impossible parce que, l'ensemble  $D$  étant fermé, nous avons  $\partial D \subset D$ .

□

### Remarque 21.93.

Bien que cela puisse paraître évident dès le début, nous ne démontrerons que dans la proposition 21.95 que  $D$  est l'enveloppe convexe de  $\Gamma$ .

CORooSXDGooJEmVef

### Corolaire 21.94.

Si  $p \in \text{Int}(D)$  alors toute droite passant par  $p$  intersecte  $\Gamma$  en exactement 2 points.

*Démonstration.* Vu que la partie  $D$  est bornée, toute droite passant par son intérieur coupe  $\partial D$  en au moins deux points (un dans chaque sens, et en utilisant le lemme de passage de douane 7.174).

Soit  $\ell$  une droite passant par  $p$  et supposons qu'elle coupe  $\Gamma$  en trois points distincts. Alors au moins deux d'entre eux sont du même côté de  $p$ . Soient  $q_1$  et  $q_2$  ces points. Nous avons donc dans l'ordre  $p \in \text{Int}(D)$ ,  $q_1 \in \partial D$  et  $q_2 \in \partial D$ .

Vu que tout  $\Gamma$  est d'un seul côté de ses tangentes, lesdites tangentes ne passent pas par l'intérieur de  $D$ . Ni  $p$  ni  $q_2$  ne sont sur  $\ell_{q_1}$ , parce que si  $q_2 \in \ell_{q_1}$  alors  $\ell_{q_1} = \ell$ , ce qui est impossible parce que  $\ell$  passe par  $p \in \text{Int}(D)$ .

Or  $p$  et  $q_2$  sont de part et d'autres de  $\ell_{q_1}$ , ce qui est impossible parce que  $\Gamma$  est d'un seul côté de cette droite. □

PROPooOORPooCXrIQi

**Proposition 21.95.**

Soit  $D$  l'ensemble défini en (21.310).

- (1)  $D = \text{Conv}(\Gamma)$  ( $\text{Conv}(\Gamma)$  désigne l'enveloppe convexe de  $\Gamma$ )
- (2)  $\partial \text{Conv}(\Gamma) = \Gamma$ .

Pour la définition d'enveloppe convexe, voir la définition 11.38.

*Démonstration.* L'ensemble  $D$  est un convexe contenant  $\Gamma$ . Donc  $\text{Conv}(\Gamma) \subset D$ . L'inclusion inverse est à prouver.

Soit  $x \in D$ . Si  $x \in \partial D$  alors  $x \in \Gamma$  (proposition 21.92) et donc  $x \in \text{Conv}(\Gamma)$ . Nous ne devons donc traiter que le cas  $x \in \text{Int}(D)$ .

Le corolaire 21.94 nous dit que toute droite passant par  $x$  coupe  $\Gamma$  en exactement deux points. Soient  $p$  et  $q$  ces points. Alors  $p, q \in \Gamma \subset \text{Conv}(\Gamma)$ . Vu que  $\text{Conv}(\Gamma)$  est convexe, tout le segment  $[p, q]$  est dans  $\text{Conv}(\Gamma)$ , et en particulier  $x \in \text{Conv}(\Gamma)$ .

Nous passons à la seconde affirmation. Nous savons que  $D = \text{Conv}(\Gamma)$ . En prenant le bord des deux côtés,  $\partial D = \partial \text{Conv}(\Gamma)$ , donc  $\Gamma = \partial \text{Conv}(\Gamma)$ .  $\square$

LEMooUEKQooWhGyKn

**Lemme 21.96** (Des tangentes parallèles[672]).

Une courbe fermée  $\gamma$  de classe  $C^1$  est convexe si et seulement si elle ne possède pas 3 tangentes parallèles distinctes.

*Démonstration.* Si le graphe  $\Gamma$  possédait trois tangentes parallèles distinctes, une serait entre les deux autres et le graphe  $\Gamma$  serait de part et d'autres de cette tangente. Dans ce cas,  $\gamma$  n'est pas convexe.

Nous montrons maintenant que si  $\gamma$  n'est pas convexe, alors elle possède trois tangentes parallèles distinctes. Pour  $p \in \Gamma$  tel que  $\Gamma$  soit des deux côtés de  $l_p$ .

Soient  $\Gamma_1$  et  $\Gamma_2$  les parties de  $\Gamma$  délimitées par  $l_p$ . Nous notons  $q_i$  le point de  $\Gamma_i$  le plus éloigné de la droite  $l_p$ , il existe parce que  $\Gamma$  est compact et que la fonction distance à  $l_p$  est continue sur  $\Gamma$ . Les points  $p, q_1$  et  $q_2$  sont distincts (sinon  $l_p$  ne couperait pas  $\Gamma$  en deux parties).

Montrons que  $l_{q_i} \parallel l_p$ . Pour cela nous choisissons un système d'axe dans lequel  $l_p \equiv y = 0$ . Dans ce système, la distance entre  $\gamma(t)$  et  $l_p$  est  $\gamma_y(t)$  et les extrémums de cette fonction ont lieu aux points  $t$  tels que  $\gamma'_y(t) = 0$ , c'est-à-dire aux points sur lesquels la tangente est parallèle à  $l_p$ .

À quel moment avons nous utilisé le fait que la courbe soit fermée? Au moment de dire que le point le plus éloigné devait vérifier  $\gamma'_y(t) = 0$ . En effet un extrémum peut ne pas vérifier cette condition si il n'est pas à l'intérieur du domaine. Dans notre cas, nous avons  $\gamma: [0, 1] \rightarrow \mathbb{R}^2$  qui est fermée :  $\gamma(0) = \gamma(1)$ . Donc en réalité nous pouvons considérer  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$  et tous les points du domaine sont intérieurs au domaine. L'extrémum doit donc vérifier la condition d'annulation de la dérivée.  $\square$

**Exemple 21.97.**

Si la courbe n'est pas fermée, alors le lemme 21.96 ne tient pas comme le montre le contre-exemple du graphe de  $f(x) = x^3 - x$ . Il est non convexe et pourtant ne présente que 2 tangentes parallèles (dans chaque direction).  $\triangle$

LEMooCSXCooIDPiKW

**Lemme 21.98.**

Soit  $\gamma$  une courbe convexe et  $\ell$  une droite qui intersecte  $\Gamma$  mais qui n'est pas tangente. Alors l'intersection entre  $\ell$  et  $\Gamma$  comprend au maximum 2 points.

*Démonstration.* Supposons que  $\ell$  coupe  $\Gamma$  en trois points  $p, q, r$  (dans cet ordre). Vu que  $\ell$  n'est pas tangente à  $\Gamma$ , la tangente  $\ell_q$  est distincte de  $\ell$  (et intersecte  $\ell$  en l'unique point  $q$ ). Les points  $p$  et  $r$  sont dans  $\Gamma$  et sont pourtant de deux côtés différents de  $\ell_q$ . Contradiction avec la convexité de  $\gamma$ .  $\square$

La proposition suivante nous dit que si deux points de  $\Gamma$  ont la même tangente, alors entre ces deux points,  $\Gamma$  est le segment de droite les joignant.

PROPOCKTZooIPcUca

**Proposition 21.99** ([1]).

Soit  $\gamma: [0, L] \rightarrow \mathbb{R}^2$  une courbe fermée simple et convexe de classe  $C^1$ . Si la tangente en  $p = \gamma(s_p)$  et la tangente en  $q = \gamma(s_q)$  sont identiques (pas seulement parallèles), alors soit

$$\gamma([s_p, s_q]) = [p, q] \quad (21.314)$$

soit

$$\gamma([s_q, L]) = [p, q] \quad (21.315)$$

*Démonstration.* Nous considérons un système d'axe dans lequel  $p = (0, 0)$ ,  $\ell_p = \ell_q \equiv y = 0$  et tel que  $\gamma_y(t) \geq 0$  pour tout  $t$ . Nous choisissons enfin un paramétrage de  $\gamma$  tel que  $p = \gamma(0)$ .

Supposons un instant que  $\gamma_y([0, s_q]) = \{0\}$ . Nous avons  $\gamma_x(0) = p_x = 0$  et  $\gamma_x(s_q) = q_x$ . Donc le théorème des valeurs intermédiaires 9.95 implique que toutes les valeurs entre 0 et  $q_x$  sont atteintes par  $\gamma_x$  entre 0 et  $s_q$ . Donc tous les points du type  $(t, 0)$  avec  $0 \leq t \leq q_x$  sont atteints par  $\gamma(t)$  avec  $0 \leq t \leq s_q$ . De plus aucun autre point ne peut être atteint parce que  $\gamma$  étant simple, elle ne peut pas faire de retour en arrière.

Nous supposons donc que  $\gamma_y$  n'est pas identiquement nulle sur  $[0, s_q]$ ; il existe donc un  $s_M$  avec  $0 < s_M < s_q$  qui maximise  $\gamma_y$  sur  $[0, s_q]$ . La tangente à  $\Gamma$  en  $\gamma(s_M)$  est horizontale. Les droites  $y = 0$  et  $y = \gamma_y(s_M)$  sont donc deux tangentes parallèles à  $\Gamma$ . Par le lemme des tangentes parallèles 21.96, il n'y a pas d'autres tangentes horizontales. Donc pour tout  $n \in \mathbb{N}$  la droite  $y = \frac{1}{n}$  n'est tangente nulle part<sup>28</sup> à  $\Gamma$ .

Par le lemme 21.98, la droite  $y = \frac{1}{n}$  ne peut intersecter  $\Gamma$  qu'en seulement deux points. Or le théorème des valeurs intermédiaires appliqué à  $\gamma_y$  sachant que  $\gamma_y(0) = \gamma_y(s_q) = 0$  et  $\gamma_y(s_M) > 0$  nous donne  $a_n, b_n$  tels que

$$0 < a_n < s_M \quad (21.316a)$$

$$s_M < b_n < s_q \quad (21.316b)$$

et  $\gamma_y(a_n) = \gamma_y(b_n) = \frac{1}{n}$ . Donc la droite  $y = \frac{1}{n}$  intersecte  $\gamma$  deux fois dans  $[0, s_q]$ . En conséquence de quoi  $\gamma_y(t) < \frac{1}{n}$  pour tout  $t \in [s_q, L]$ . Cela étant valable pour tout  $n$  nous avons  $\gamma_y([s_q, L]) = \{0\}$  et nous sommes ramenés essentiellement au premier cas.  $\square$

PROPOCKTZooIPcUca

**Proposition 21.100.**

Soit une courbe fermée simple  $\gamma$  de classe  $C^1$  et une droite  $\ell$ . Alors il existe au moins deux points distincts  $q_1, q_2$  tels que  $\ell_{q_1} \parallel \ell_{q_2} \parallel \ell$  avec  $\ell_{q_1} \neq \ell_{q_2}$ .

*Démonstration.* Nous considérons un système d'axes dans lequel la droite  $\ell$  est la droite  $y = 0$ <sup>29</sup>. Nous considérons la fonction

$$\begin{aligned} r: \gamma &\rightarrow \mathbb{R} \\ p &\mapsto p_y. \end{aligned} \quad (21.317)$$

En tant que fonction continue sur un compact, elle possède un minimum et un maximum. Les extrémums arrivent en  $\gamma(s)$  avec  $\gamma'_y(s) = 0$ , ce qui signifie que les tangentes aux extrémums sont parallèles à  $\ell$ . Vu que le maximum et le minimum ne peuvent pas être égaux (sinon la courbe serait horizontale et pas simple), les tangentes en ces points sont distinctes.  $\square$

**Corolaire 21.101.**

Soit une courbe convexe fermée simple  $\gamma$  de classe  $C^2$ . L'ensemble  $\text{Conv}(\Gamma)$  est compact.

*Démonstration.* Nous savons que  $\text{Conv}(\Gamma)$  n'est autre que  $D$  par la proposition 21.95. Nous savons déjà que  $D$  est fermé. Il nous suffit donc de prouver qu'il est borné (théorème de Borel-Lebesgue 9.24). Nous considérons deux droites perpendiculaires et les 4 tangentes correspondantes par la proposition 21.100. Vu que  $\Gamma$  est d'un seul côté de chacune de ces tangentes, elle est contenue dans le rectangle délimité par ces 4 droites.  $\square$

28. À part  $n = 0$  et si par manque de chance,  $\gamma_y(s_M)$  est un nombre de la forme  $1/n$ .

29. Pour bien faire, il faut faire la démonstration en deux parties : une qui montre le cas particulier  $\ell \equiv y = 0$ , et un second qui en déduit rigoureusement le cas général.

## 21.13.4 Courbure et convexité

SUBSECooNJOLooYuGRjA

À propos de paramétrage normal.

LEMooHMFSooflhanD

**Lemme 21.102.**

Soit  $\gamma: [0, L] \rightarrow \mathbb{R}^2$  une courbe simple, fermée de classe  $C^2$  en paramétrage normal. Nous considérons sa fonction d'angle  $\theta: [0, L] \rightarrow \mathbb{R}$ .

Alors l'application

$$\begin{aligned} \sigma: [0, L] &\rightarrow S^1 \\ t &\mapsto \begin{pmatrix} \cos(\theta(t)) \\ \sin(\theta(t)) \end{pmatrix} \end{aligned} \quad (21.318)$$

est surjective.

*Démonstration.* Le théorème 21.88 nous dit que si  $\theta(0) = a$  alors  $\theta(2\pi)$  ne peut valoir que  $a + 2\pi$  ou  $a - 2\pi$ . Le théorème des valeurs intermédiaires nous dit alors que  $\theta$  prend toutes les valeurs entre  $a$  et  $a + 2\pi$  ou  $a - 2\pi$ .  $\square$

PROPooWXUKooPOtPdJ

**Proposition 21.103** ([672]).

Une courbe fermée simple de classe  $C^2$  est convexe si et seulement si sa courbure est soit toujours positive, soit toujours négative.

*Démonstration.* Nous considérons la courbe  $\gamma$  munie d'un paramétrage de vitesse 1, c'est-à-dire avec  $\|\gamma'(t)\| = 1$  pour tout  $t$ . Si  $\theta$  est sa fonction d'angle<sup>30</sup>, alors nous avons  $\theta' = \kappa$  par le lemme 21.80. Donc la fonction  $\theta$  est monotone si et seulement si la courbure ne change pas de signe. Nous allons montrer que  $\theta$  est monotone si et seulement si  $\gamma$  est convexe.

(1)  $\Rightarrow$

Nous supposons que  $\theta$  est monotone et  $\gamma$  non convexe. Soient  $p \in \Gamma$  tel que  $\Gamma$  soit des deux côtés de  $\ell_p$ , et soient les points  $q_1, q_2$  donnés par le lemme des tangentes parallèles 21.96 tels que  $\ell_p \parallel \ell_{q_1} \parallel \ell_{q_2}$ . Parmi les vecteurs tangents en  $p, q_1$  et  $q_2$ , deux au moins ont la même direction; supposons que ce sont  $q_1$  et  $q_2$ . C'est-à-dire que si  $p = \gamma(s_0)$ ,  $q_1 = \gamma(s_1)$  et  $q_2 = \gamma(s_2)$  alors nous avons  $\gamma'(s_1) = \gamma'(s_2)$  et donc aussi

$$\theta(s_1) = \theta(s_2) + 2n\pi \quad (21.319)$$

pour un certain  $n$ . Mais  $\theta$  est monotone et la différence entre sa première et sa dernière valeur doit valoir  $2\pi$  ou  $-2\pi$  par le théorème 21.88. Donc  $n$  ne peut valoir que  $-1, 0$  et  $1$ .

Si  $n = 0$  alors  $\theta$  est constante sur  $[s_1, s_2]$ . Si  $n = 1$  alors  $\theta(s_1) = \theta(s_2) + 2\pi$  alors que sur toute la courbe,  $\theta$  ne peut faire que  $2\pi$ . Donc  $\theta$  est constant sur  $[0, s_1]$  et sur  $[s_2, L]$  (où  $L$  est le bord de le paramétrage). Si  $n = -1$ , même conclusion.

Dans tous les cas,  $\Gamma$  contient une ligne droite, soit de  $q_1$  à  $q_2$ , soit de  $q_2$  à  $q_1$ . Et dans ces cas nous avons  $\ell_{q_1} = \ell_{q_2}$ , ce qui est contraire à la construction de  $q_i$ .

Nous concluons que  $\gamma$  est convexe.

(2)  $\Leftarrow$

Nous supposons que  $\gamma$  est convexe, mais que  $\theta$  n'est pas monotone. Il existe donc  $s_1 < s_0 < s_2$  tels que

$$\theta(s_1) = \theta(s_2) \neq \theta(s_0). \quad (21.320)$$

Et vu le lemme 21.102, il existe  $s_3$  tel que  $\gamma'(s_3) = -\gamma'(s_1)$ .

Donc en  $s_1, s_2$  et  $s_3$  nous avons trois tangentes parallèles. La proposition 21.96 est alors formelle,  $\gamma$  étant convexe, deux de ces tangentes doivent être identiques.

La proposition 21.99 dit qu'entre deux points dont les tangentes sont identiques, la courbe doit être un segment de droite. Or sur un segment de droite,  $\kappa = 0$  et  $\theta$  est constante.

30. Définition 21.79.



- La partie  $\gamma([s_1, s_2])$  ne peut pas être droite parce que nous avons supposé l'existence d'un  $s_0 \in ]s_1, s_2[$  tel que  $\theta(s_1) = \theta(s_2) \neq \theta(s_0)$ .
- La partie  $\gamma([s_1, s_3])$  ne peut pas être droite parce que  $\theta(s_3) \neq \theta(s_1)$ .
- La partie  $\gamma([s_2, s_3])$  ne peut pas être droite parce que  $\theta(s_2) \neq \theta(s_1)$ .

Nous sommes donc devant une contradiction.

Nous en concluons que  $\theta$  doit être monotone.

□

### 21.13.5 Théorème des quatre sommets

LEMooELIRooNDVXPh

#### Lemme 21.104.

Soit une droite  $\ell$  du plan. Il existe  $a, c \in \mathbb{R}^2$  avec  $c \neq 0$  tels que  $z \in \ell$  si et seulement si  $(z - a) \cdot c = 0$ .

*Démonstration.* Une droite est paramétrée par  $\gamma(t) = p + tq$ . En posant  $a = p$  et  $c = Jq$  nous avons la réponse. En effet nous allons montrer qu'avec ces valeurs de  $a$  et  $c$ , nous avons  $z \in \Gamma$  si et seulement si  $(z - a) \cdot c = 0$ .

D'abord un point de  $\gamma$  est de la forme  $z = \gamma(t) = p + tq$ . Nous avons :

$$(\gamma(t) - a) \cdot c = (\gamma(t) - p) \cdot Jq = tq \cdot Jq = 0. \quad (21.321)$$

Et dans l'autre sens, si  $(z - a) \cdot c = 0$  nous devons prouver que  $z \in \Gamma$ . Nous avons :  $(z - p) \cdot Jq = 0$ , ce qui fait que  $z - p$  est un multiple de  $q$ . Autrement dit :  $z - p = \lambda q$  ou encore  $z = \alpha q + p$ , qui est sur la droite  $\Gamma$ . □

#### Définition 21.105.

Un **sommet** d'une courbe est un point d'extrémum local de la courbure.

THOooFRBBooWKZcfY

#### Théorème 21.106 (Théorème des quatre sommets[126, 603]).

Soit un arc paramétrique  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$  fermé, simple et convexe<sup>31</sup> de classe  $C^3$  et  $T$ -périodique.

Alors  $\gamma$  possède au moins 4 points critiques sur chaque période.

*Démonstration.* Nous supposons que le paramétrage de  $\gamma$  soit normal.

Si la courbure  $\kappa$  est constante sur une partie ouverte de la (du paramétrage de la) courbe, alors tous les points de cette partie sont des sommets et le théorème est fait. Nous supposons que  $\kappa$  n'est pas constante et en particulier que  $\Gamma$  ne contient ni bouts de droites ni bouts de cercles (théorème 21.76).

La fonction  $\kappa$  étant de classe  $C^1$  sur le compact  $\Gamma$ , elle admet au moins un maximum et un minimum distincts. Vu que ces points sont intérieurs, ils correspondent au changement de signe de  $\kappa'$ . Soient  $p$  et  $q$  ces points. Pour la simplicité nous supposons que  $\gamma$  est paramétré de telle sorte que  $\gamma(0) = p$ , et  $q = \gamma(s_q)$  avec  $0 < s_q < T$ .

Nous supposons que  $p$  et  $q$  sont les seuls points de changement de signe de  $\kappa'$ .

Soit  $\ell$  la droite passant par  $p$  et  $q$ . Tous les points du segment  $[p, q]$  (qui sont dans  $\text{Conv}(\Gamma)$ ) ne peuvent pas être sur  $\Gamma$  (sinon nous aurions un morceau de droite). Donc certains points sont dans l'intérieur de  $\text{Conv}(\Gamma)$ . Donc la droite  $\ell$  passe par l'intérieur de  $\text{Conv}(\Gamma)$  et le corolaire 21.94 nous dit que la droite  $\ell$  ne coupe  $\Gamma$  en seulement deux points.

Par conséquent, les ensembles  $\gamma([0, s_q])$  et  $\gamma([s_q, T])$  sont de part et d'autre de  $\Gamma$ . Vu qu'en ces points,  $\kappa'$  change de signe et qu'il ne change de signe en aucun autre points, la fonction  $\kappa'$  est positive d'un côté de  $\ell$  et négative de l'autre.

D'autre part par le lemme 21.104, il existe  $a \in \mathbb{R}^2$  et  $c \neq 0$  tels que  $z \in \ell$  si et seulement si  $(z - a) \cdot c = 0$ . La fonction  $z \mapsto (z - a) \cdot c$  est donc positive d'un côté de  $\ell$  et négative de l'autre.

31. Par la proposition 21.103 nous pouvons aussi bien demander à la courbure d'être toujours strictement positive, comme le fait [126].

En résumé les fonctions

$$s \mapsto \kappa'(s) \quad (21.322a)$$

$$s \mapsto (\gamma(s) - a) \cdot c \quad (21.322b)$$

changent de signe en même temps et le produit a donc un signe constant. Ce produit n'est de plus pas nul parce que  $\kappa'$  n'est nul sur aucun intervalle (sinon  $\kappa$  y serait constant et  $\Gamma$  un segment de droite) et  $(\gamma(s) - a) \cdot c$  ne s'annule pour aucun  $s$  sauf ceux qui correspondent à  $p$  et  $q$ .

Nous avons donc

$$0 \neq \int_0^T \kappa'(s)(\gamma(s) - a) \cdot c \, ds \quad (21.323a)$$

$$= \underbrace{\left[ (\gamma(s) - a) \cdot c \kappa(s) \right]_0^T}_{A=0} - \int_0^T \kappa(s) \gamma'(s) \cdot c \, ds \quad (21.323b)$$

$$= - \int_0^T \kappa(s) (\gamma'(s) \cdot c) \, ds \quad (21.323c)$$

$$= \int_0^T J \gamma''(s) \cdot c \, ds \quad (21.323d)$$

$$= J \int_0^T \gamma''(s) \cdot c \, ds \quad (21.323e)$$

$$= J [\gamma'(s) \cdot c]_0^T \quad (21.323f)$$

$$= 0. \quad (21.323g)$$

Justifications :

- L'expression  $A$  est nulle parce que les valeurs en 0 et en  $T$  sont identiques.
- Nous utilisons le lemme 21.75 pour faire  $-\kappa(s)\gamma'(s) = J\gamma''(s)$ .

Le tout est une contradiction de la forme  $0 \neq a = 0$ .

Nous avons donc au moins un troisième point de changement de signe de  $\kappa'$ . Vu que la courbe est périodique, il en faut un nombre pair et donc un quatrième.  $\square$

L'exemple de l'ellipse montre qu'il n'y a pas lieu de chercher d'autres extrémums de  $\kappa$  à part les 4 déjà trouvés.

### Exemple 21.107.

Nous trouvons les sommets de l'ellipse.

$$\gamma(t) = (a \cos(t), b \sin(t)) \quad (21.324a)$$

$$\gamma'(t) = (-a \sin(t), b \cos(t)) \quad (21.324b)$$

$$\gamma''(t) = -(a \cos(t), b \sin(t)) \quad (21.324c)$$

$$(21.324d)$$

La courbure est

$$\kappa(t) = \frac{\gamma'' \cdot J\gamma'}{\|\gamma'\|^3} \quad (21.325a)$$

$$= \frac{-1}{[a^2 \sin^2(t) + b^2 \cos^2(t)]^{3/2}} \begin{pmatrix} a \cos(t) \\ b \sin(t) \end{pmatrix} \cdot \begin{pmatrix} -b \cos(t) \\ -a \sin(t) \end{pmatrix} \quad (21.325b)$$

$$= \frac{ab}{[a^2 \sin^2(t) + b^2 \cos^2(t)]^{3/2}}. \quad (21.325c)$$

Vu que  $ab > 0$ , les extrémums de cela sont ceux du dénominateur et il suffit donc d'étudier les extrémums de

$$f(t) = a^2 \sin^2(t) + b^2 \cos^2(t). \quad (21.326)$$



Nous avons

$$f'(t) = 2(a^2 - b^2) \cos(t) \sin(t), \quad (21.327)$$

fonction qui s'annule effectivement 4 fois sur une période. Deux maximums et deux minima.  $\triangle$

### 21.13.6 Théorème d'Urysohn

LEMooWRLRooCdubfZ

#### Lemme 21.108.

Soit un espace topologique  $X$ . Soit une partie  $D$  dense de  $[0, 1]$ . Nous considérons des ouverts  $\{U_s\}_{s \in D}$  de  $X$  indexés par  $D$ . Nous supposons que si  $s_1 < s_2$  alors  $\bar{U}_{s_1} \subset U_{s_2}$ . Enfin pour tout  $x \in X$  nous posons

$$A_x = \{s \in D \text{ tel que } x \in U_s\}. \quad (21.328)$$

Alors l'application

$$f: X \rightarrow [0, 1]$$

$$x \mapsto \begin{cases} \inf(A_x) & \text{si } x \in \bigcup_{s \in D} U_s \\ 1 & \text{sinon.} \end{cases} \quad (21.329)$$

est continue.

*Démonstration.* Dans la suite nous prenons les conventions suivantes :

$$\bigcup_{s \in \emptyset} U_s = \emptyset \quad (21.330a)$$

$$\bigcap_{s \in \emptyset} U_s = X \quad (21.330b)$$

$$(21.330c)$$

Soit  $t \in [0, 1]$ .

- (1)  $\{x \text{ tel que } f(x) < t\} = \bigcup_{s < t} U_s$  D'abord nous considérons  $x \in X$  tel que  $f(x) < t$ , et nous prouvons que  $x \in \bigcup_{s < t} U_s$ . Vu que  $f(x) = \inf(A_x) < t \leq 1$ , il existe  $s_0 < t$  tel que  $s_0 \in A_x$ . Pour ce  $s_0$  nous avons  $x \in U_{s_0}$ , et donc

$$x \in U_{s_0} \subset \bigcup_{\substack{s \in D \\ s < t}} U_s. \quad (21.331)$$

Dans l'autre sens, soient  $s_0 < t$  et  $x \in U_{s_0}$ . Nous avons  $s_0 \in A_x$  et donc  $f(x) \leq s_0 < t$ .

- (2)  $\{x \text{ tel que } f(x) \leq t\} = \bigcup_{s \geq t} \bar{U}_s$  Si  $t = 1$ , alors  $\{x \text{ tel que } f(x) \leq t\} = X$ , et  $\bigcap_{s > t} \bar{U}_s = \bigcup_{s \in \emptyset} \bar{U}_s = X$ . Supposons que  $t < 1$ .

Soit  $x \in X$  tel que  $f(x) \leq t$ . Vu que  $t < 1$  et que  $D$  est dense dans  $[0, 1]$ , il existe  $s \in D$  tel que

$$s > t \geq f(x) = \inf(A_x). \quad (21.332)$$

Donc il existe  $a \in A_x$  tel que  $a \in ]\inf(A_x), s[$ . Vu que  $a \in A_x$ , nous avons  $x \in U_a$ , et vu que  $a < s$ , nous avons  $\bar{U}_a \subset U_s$ . Donc

$$x \in U_a \subset \bar{U}_a \subset U_s. \quad (21.333)$$

Donc  $x \in U_s$  pour tout  $s > t$  (dans  $D$ ), et donc  $x \in \bigcap_{s > t} \bar{U}_s$ . Nous avons donc déjà prouvé que

$$\{x \text{ tel que } f(x) \leq t\} \subset \bigcap_{\substack{s \in D \\ s > t}} U_s. \quad (21.334)$$

Nous faisons à présent l'inclusion dans l'autre sens. Soit  $x \in \bigcap_{\substack{s \in D \\ s > t}} \bar{U}_s$ , et montrons que  $f(x) \leq t$ . Nous avons déjà traité le cas  $t = 1$ ; nous continuons à supposer que  $t < 1$ .

Soit  $\epsilon > 0$  assez petit pour avoir  $t + \epsilon < 1$ . Vu que  $D$  est dense dans  $[0, 1]$ , il existe  $s_1, s_2 \in D$  tels que

$$t < s_1 < s_2 < t + \epsilon. \quad (21.335)$$

Par hypothèse  $x \in \bar{U}_{s_1}$  et donc

$$x \in \bar{U}_{s_1} \subset U_{s_2}. \quad (21.336)$$

Cela prouve que  $s_2 \in A_x$ , et nous avons alors

$$f(x) = \inf(A_x) \leq s_2 < t + \epsilon. \quad (21.337)$$

Étant donné que  $\epsilon$  est arbitrairement petit, nous avons  $f(x) \leq t$ .

- (3)  $\{x \text{ tel que } f(x) < t\}$  est ouvert pour tout  $t \in \mathbb{R}$  Nous commençons par  $t \in [0, 1]$ . Dans ce cas nous avons vu que

$$\{x \text{ tel que } f(x) < t\} = \bigcup_{\substack{s \in D \\ s < t}} U_s \quad (21.338)$$

qui est une union d'ouverts, et donc qui est ouvert.

Si  $t < 0$ , alors

$$\bigcup_{s < t} U_s = \bigcup_{s \in \emptyset} U_s = \emptyset, \quad (21.339)$$

est ouvert. Si  $t > 1$ , alors  $\{x \text{ tel que } f(x) < t\} = X$  parce que  $f$  prend ses valeurs dans  $[0, 1]$ .

- (4)  $\{x \text{ tel que } f(x) \leq t\}$  est fermé pour tout  $t \in \mathbb{R}$  Si  $t \in [0, 1]$ , alors nous savons déjà que

$$\{x \text{ tel que } f(x) \leq t\} = \bigcup_{s > t} \bar{U}_s \quad (21.340)$$

qui est fermé comme intersection de fermés (lemme 7.6(1)). Vu que  $f$  prend ses valeurs dans  $[0, 1]$ , nous avons  $\{x \text{ tel que } f(x) < t\} = \emptyset$  lorsque  $t < 0$  et  $\{x \text{ tel que } f(x) < t\} = X$  lorsque  $t > 1$ .

$f$  est continue Soient  $a, b \in \mathbb{R}$  avec  $a < b$ . Nous avons,

$$f^{-1}(]a, b[) = \{f < b\} \setminus \{f \leq a\} = \{f < b\} \cap \{f \leq a\}^c. \quad (21.341)$$

Mais nous savons que  $\{f < b\}$  est ouvert et que  $\{f \leq a\}$  est fermé. Donc  $f^{-1}(]a, b[)$  est ouvert comme intersection d'ouverts.

La proposition 7.188 dit (entre autres) que les  $]a, b[$  forment une base de la topologie de  $\mathbb{R}$ . La proposition 7.47 dit que cela suffit pour conclure à la continuité de  $f$ .

□

THOooKYYEooLFCNpg

**Théorème 21.109** (Théorème d'Urysohn[234]).

Soit un espace topologique normal  $X$ . Soient  $A$  et  $B$ , des fermés disjoints de  $X$ . Il existe une application continue  $f: X \rightarrow [0, 1]$  telle que  $f|_A = 0$  et  $f|_B = 1$ .

*Démonstration.* Pour chaque  $m \in \mathbb{N}$  nous considérons  $D_m = \{k/2^m\}_{k=0, \dots, 2^m}$ , et posons  $D = \bigcup_{m=0}^{\infty} D_m$ .

- (1)  $D$  est dense dans  $[0, 1]$  Soient  $x \in [0, 1]$  et  $\epsilon > 0$ . Considérons  $m$  tel que  $1/2^m < \epsilon$ . Dans ce cas, il existe  $s \in D_m$  tel que  $|s - x| < \frac{1}{2^m} < \epsilon$ .
- (2) Construction des  $U_s$  Nous allons maintenant construire un ouvert  $U_s$  pour chaque  $s \in D$ . Et nous allons le faire de telle sorte d'avoir  $s_1 < s_2 \Rightarrow \bar{U}_{s_1} \subset U_{s_2}$ .
- (3) Pour  $s \in D_0$  Nous commençons avec les éléments de  $D_0 = \{0, 1\}$ . D'abord nous posons

$$U_1 = B^c. \quad (21.342)$$

Vu que  $B$  est fermé, c'est un ouvert.

Pour construire  $U_0$ , nous remarquons que  $A$  est un fermé disjoint de  $B$  dans l'espace normal  $X$ . Il existe donc des ouverts  $U_0$  et  $U'_0$  qui séparent  $A$  et  $B$ , c'est-à-dire tels que  $A \subset U_0$ ,  $B \subset U'_0$  et  $U_0 \cap U'_0 = \emptyset$ .

Nous avons  $\bar{U}_0 \subset B^c$ . En effet si  $a \in \bar{U}_0 \cap B$ , alors  $a \in B$  et donc  $U'_0$  est un voisinage ouvert de  $a$ . Ce voisinage n'intersectant pas  $U_0$ ,  $a$  n'est pas dans la fermeture de  $U_0$ .

Nous avons donc bien la relation

$$\bar{U}_0 \subset B^c = U_1. \quad (21.343)$$

- (4) **Réurrence** Nous avons défini  $U_s$  pour  $s \in D_0$ . Pour la récurrence, nous supposons que  $U_s$  est défini pour  $s \in D_{m-1}$ , et qu'ils vérifient la condition. Nous allons construire les  $U_s$  pour  $s \in D_m$ .

Remarquons que  $D_{m-1} \subset D_m$ . Plus précisément, les éléments de  $D_m$  qui ne sont pas dans  $D_{m-1}$  sont ceux de la forme  $k/2^m$  avec  $k$  impair.

Soit donc  $k$  impair et considérons  $s = k/2^m$ . Par hypothèse de récurrence nous avons déjà

$$\bar{U}_{(k-1)/2^m} \subset U_{(k+1)/2^m}. \quad (21.344)$$

Le lemme 7.93 nous permet de considérer un ouvert  $U$  tel que

$$\bar{U}_{(k-1)/2^m} \subset U \subset \bar{U} \subset U_{(k+1)/2^m}. \quad (21.345)$$

Nous posons  $U_{k/2^m} = U$  et nous prouvons que  $\bar{U}_{s_1} \subset U_{s_2}$  dès que  $s_1 < s_2$  dans  $D_m$ .

Soient  $s_1 = k/2^m$  et  $s_2 = l/2^m$  avec  $k < l$ . Il y a quatre possibilités suivant la parité de  $k$  et  $l$ .

- (4a) **Si  $k$  et  $l$  sont pairs** Alors  $s_1$  et  $s_2$  sont dans  $D_{m-1}$  et l'hypothèse de récurrence dit  $\bar{U}_{s_1} \subset U_{s_2}$ .
- (4b) **Si  $k$  est impair et  $l$  est pair** Nous notons  $k = 2k' + 1$  et  $l = 2l'$ . Vu que  $k < l$ , nous avons  $k' + \frac{1}{2} < l'$ . Mais comme  $k'$  et  $l'$  sont entiers, nous avons  $k' + 1 \leq l'$ . Et nous sommes partis pour le calcul :

$$\bar{U}_{s_1} = \bar{U}_{k/2^m} \quad (21.346a)$$

$$\subset U_{(k+1)/2^m} \quad \text{construction pour } k \text{ pair} \quad (21.346b)$$

$$= U_{(k'+1)/2^{m-1}} \quad (21.346c)$$

$$\subset U_{l'/2^{m-1}} \quad \text{hyp. rec. dans } D_{m-1} \quad (21.346d)$$

$$= U_{l/2^m} \quad (21.346e)$$

$$= U_{s_2}. \quad (21.346f)$$

- (4c) **Si  $k$  est pair et  $l$  est impair** Laisseé aux bons soins du lecteur. J'ai pas vérifié, mais [234] dit que c'est similaire. Écrivez-moi si vous voyez un problème.

- (4d) **Si  $k$  et  $l$  sont impairs** Laisseé aux bons soins du lecteur. J'ai pas vérifié, mais [234] dit que c'est similaire. Écrivez-moi si vous voyez un problème.

- (5) **Résumé** Nous avons posé  $D_m = \{k/2^m\}_{k=0,\dots,2^m}$ , et  $D = \bigcup_{m=0}^{\infty} D_m$ . Nous avons prouvé que  $D$  est dense dans  $[0, 1]$ , et nous avons construit, pour chaque  $s \in D$  un ouvert  $U_s$  de  $X$ , de telle façon que

$$\bar{U}_{s_1} \subset U_{s_2} \quad (21.347)$$

pour tout  $s_1 < s_2$ . Bref, le lemme 21.108 est applicable. Nous posons donc

$$f: X \rightarrow [0, 1] \quad (21.348)$$

$$x \mapsto \begin{cases} \inf(A_x) & \text{si } x \in \bigcup_{s \in D} U_s \\ 1 & \text{sinon.} \end{cases}$$

avec  $A_x = \{s \in D \text{ tel que } x \in U_s\}$ .

(6) Ce  $f$  est ce qu'il nous fait Si  $x \in B$ , alors  $x$  n'est pas dans  $U_1$  et donc n'est dans aucun des  $U_s$ . Donc  $f(x) = 1$ .

Si  $x \in A$ , alors nous avons  $x \in U_0$  (parce que  $A \subset U_0$ ). En particulier  $0 \in A_x$ , et donc  $f(x) = \inf(A_x) = 0$ .

□

**21.110.**

Les fractions de la forme  $a/2^b$  sont nommées **fractions dyadiques** et nous avons déjà vu qu'elles sont denses dans  $\mathbb{R}$  dans le corolaire 1.549.



# Chapitre 22

## Géométrie hyperbolique

### 22.1 Inversion

LEMooNLYKooKBMdDk

#### Lemme-Définition 22.1.

Soit un cercle  $C$  et un point  $A$  du plan  $\mathbb{R}^2$ . Soit une droite passant par  $A$  et coupant  $C$  en deux points  $P$  et  $P'$  (pas spécialement distincts). Alors le nombre

$$\|AP\|\|AP'\| \quad (22.1)$$

ne dépend pas du choix de la droite et est nommé la **puissance** du point  $A$  par rapport au cercle  $C$ .

#### 22.1.1 Cercles perpendiculaires

##### Définition 22.2.

Deux cercles sont **perpendiculaires** lorsque leurs tangentes aux points d'intersection sont perpendiculaires.

LEMooWMGOooAieUjD

##### Lemme 22.3 ([1]).

Soient deux cercles perpendiculaires  $C_1$  et  $C_2$ . Alors

ITEMooJVYYooSriSdA

(1) le centre de  $C_1$  est hors de  $C_2$ .

ITEMooTQATooNWvllF

(2) Si  $\ell$  est une droite passant par le centre de  $C_1$  (nommé  $O$ ) et si  $\ell$  coupe  $C_2$  en les deux points  $P$  et  $P'$ , alors  $P$  et  $P'$  sont situés du même côté de  $O$ .

*Démonstration.* Nous nommons  $O_1$  le centre de  $C_1$  ainsi que  $Q$  et  $Q'$  les points d'intersection de  $C_1$  avec  $C_2$ . Si  $\ell_Q$  et  $\ell_{Q'}$  sont les tangentes à  $C_1$  en  $Q$  et  $Q'$ , alors ce sont des rayons de  $C_2$  (parce que les cercles sont perpendiculaires). Par conséquent le centre  $O_2$  de  $C_2$  est le point d'intersection  $O_2 = \ell_Q \cap \ell_{Q'}$ .

Le triangle  $O_1O_2Q$  est rectangle en  $Q$  et donc  $\|O_1O_2\| > \|QO_2\|$ . Or le nombre  $\|QO_2\|$  est le rayon de  $C_2$ , donc  $O_1$  est en dehors de  $C_2$ .

Ceci achève de prouver le point (1); nous démontrons le point (2). Les points  $Q$  et  $Q'$  sont sur le cercle  $C_2$ , donc tous les points du segment  $[QQ']$  sont dans le cercle. Or le centre de  $C_1$  doit être en dehors de  $C_2$ ; il ne peut donc pas être dans le segment  $[QQ']$ , ce qui prouve que  $Q$  et  $Q'$  ne sont pas de part et d'autre de  $O$  sur la droite  $(OQQ')$ .  $\square$

PROPooYSVYooOFKxib

##### Proposition 22.4 ([1]).

Soit un cercle  $C_1$  de centre  $O$  et de rayon  $R$ .

ITEMooWYIJooAAmXUI

(1) Un cercle  $C_2$  est perpendiculaire à  $C_1$  si et seulement si il existe une droite  $\ell$  passant par  $O$  telle que les points d'intersection  $\{P, P'\} = \ell \cap C_2$  soient situés du même côté de  $O$  et vérifient

$$\|OP\|\|OP'\| = R^2. \quad (22.2)$$

ITEMooYKZOooYZKyhF

(2) Dans ce cas, toutes les droites passant par  $O$  et coupant  $C_2$  en deux points  $P, P'$  vérifient le fait que  $P$  et  $P'$  soient du même côté de  $O$  et  $\|OP\|\|OP'\| = R^2$ .

*Démonstration.* Nous commençons par prouver le point (2). Le fait que les points  $P$  et  $P'$  soient du même côté de  $O$  est le lemme 22.3(2). Pour la relation sur les distances, soit  $Q \in C_1 \cap C_2$ . Vu que  $C_1$  et  $C_2$  sont perpendiculaires, la droite  $(OQ)$  ne coupe  $C_2$  qu'en ce point, et la puissance de  $O$  par rapport au cercle  $C_2$  est  $\|OQ\|^2 = R^2$ .

La même puissance peut être calculée via la droite  $\ell$  :

$$\|OP\|\|OP'\|. \quad (22.3)$$

Donc  $\|OP\|\|OP'\| = R^2$ .

Soit un cercle  $C_2$  passant par  $P$  et  $P'$ . Notons que  $P$  et  $P'$  ne sont pas sur  $C_1$  parce qu'ils ne pourraient pas être alignés avec  $O$ . De plus l'un est à l'intérieur de  $C_1$  et l'autre à l'extérieur de  $C_1$ . Les cercles  $C_1$  et  $C_2$  possèdent donc deux points distincts d'intersections. La puissance de  $O$  par rapport à  $C_2$  est :

$$\|OP\|\|OP'\| = R^2 \quad (22.4)$$

parce que  $(OP)$  est une droite coupant  $C_2$  en les points  $P$  et  $P'$ .

Soit  $Q$  un point d'intersection de  $C_1$  et  $C_2$ , et  $Q'$  l'autre point d'intersection de  $C_2$  avec la droite  $(OQ)$ . La puissance de  $O$  par rapport à  $C_2$  peut également être calculée à partir de cette droite (lemme 22.1) et nous avons

$$\|OQ\|\|OQ'\| = R^2, \quad (22.5)$$

mais  $Q \in C_1$ , donc  $\|OQ\| = R$  et partant  $\|OQ'\| = R$ . Nous en déduisons que  $Q' \in C_1$  également. Or  $Q'$  ne peut pas être l'autre point d'intersection de  $C_1$  avec  $C_2$  (sinon  $O, Q, Q'$  ne seraient pas alignés). Donc  $Q = Q'$  et nous déduisons que la droite  $(OQ)$  est tangente à  $C_2$ .

Le rayon de  $C_1$  est tangent à  $C_2$ . Cela signifie que  $C_1$  est perpendiculaire à  $C_2$ .  $\square$

### 22.1.2 Inversion

PROPDEFooVLIWooQgpLQa

**Proposition-Définition 22.5** ([605]).

Soit un cercle  $C$  de centre  $O$  dans  $\mathbb{R}^2$ . Il existe une unique application

$$i_C: \mathbb{R}^2 \setminus \{O\} \rightarrow \mathbb{R}^2 \setminus \{O\} \quad (22.6)$$

telle que

(1)  $i_C(x) = x$  pour tout  $x \in C$

ITEMooXLZCooEGAxHu

(2)  $i_C$  échange l'intérieur et l'extérieur de  $C$ .

ITEMooCPPUooDJLzSk

(3)  $i_C$  laisse invariants les droites et les cercles orthogonales à  $C$ .

Cette application est l'**inversion** de cercle  $C$ .

*Démonstration.* Soit  $P$  à l'intérieur de  $C$ , mais différent de  $O$ . Nous notons  $\ell$  la droite  $(OP)$  et nous considérons un cercle  $C_2$  passant par  $P$  et perpendiculaire à  $C$  (existence par la proposition 22.4). Nous avons  $\ell \perp C$  (parce que  $\ell$  est un rayon) et  $C_2 \perp C$ . Donc  $i_C(\ell) = \ell$  et  $i_C(C_2) = C_2$  par l'exigence (3). Mais comme  $P \in \ell \cap C_2$  nous avons aussi

$$i_C(P) \in \ell \cap C_2. \quad (22.7)$$

Mais  $\ell$  et  $C_2$  se coupent en exactement deux points. Vu que  $i_C(P)$  doit être hors de  $C$  (exigence (2)), avoir  $i_C(P) = P$  est impossible. Nous en concluons que  $i_C(P)$  doit être l'autre intersection.

Nous avons prouvé que les conditions (2) et (3) fixent l'image d'un point situé dans l'intérieur de  $C$ .

Si  $P$  est extérieur au cercle  $C$ , la même procédure fonctionne : nous considérons la droite  $\ell = (OP)$  et un cercle  $C_2$  perpendiculaire à  $C$  et passant par  $P$ . Encore une fois, ces deux objets sont fixés par  $i_C$ , et vu que  $i_C(P)$  doit être à l'intérieur de  $C$ , il est fixé.

L'unicité est montrée.

En ce qui concerne l'existence, si  $P \neq O$ , la procédure suivante donne  $P'$  :

- Soit la droite  $\ell = (OP)$ .
- Soit un cercle  $C_2$  perpendiculaire à  $C$  et passant par  $P$ .
- Le point  $P'$  est le point de l'intersection  $C_2 \cap \ell$  qui n'est pas  $P$ .

Il est aisé de vérifier que poser  $i_C(P) = P'$  donne une application qui vérifie toutes les propriétés.  $\square$

NORMooUBFQooCzXRJj

## 22.6.

Nous recopions la construction de l'inversion d'un point par rapport à un cercle. Si  $C_1$  est un cercle de centre  $O$  et  $P$  est un point différent de  $O$ , alors la procédure suivante construit  $P' = i_{C_1}(P)$  :

- Soit la droite  $\ell = (OP)$ .
- Soit un cercle  $C_2$  perpendiculaire à  $C_1$  et passant par  $P$ .
- Le point  $P'$  est le point de l'intersection  $C_2 \cap \ell$  qui n'est pas  $P$ .

Que se passe-t-il si  $\ell$  et  $C_2$  n'ont qu'une seule intersection ? Alors la droite  $\ell = (OP)$  est tangente à  $C_2$ . Or de  $O$  il n'existent que deux droites tangentes à  $C_2$ , et ce sont les rayons passant par les intersections parce que les cercles sont perpendiculaires. En d'autres mots, cette situation se présente lorsque  $P$  est sur le cercle  $C_1$ . Dans ce cas,  $i_{C_1}(P) = P$ .

## Remarque 22.7.

Lorsque nous disons qu'une inversion « conserve les droites passant par  $O$  », il y a pour sous-entendu que nous considérons la droite privée du point  $O$ , parce que de toutes façons, l'inversion n'est pas définie sur  $O$ .

Nous allons résoudre cet intéressant problème en 23.6.1, en ajoutant le point  $\infty$  à toutes les droites.

## Corolaire 22.8.

L'inversion  $i_C$  est une involution ( $i_C^2 = \text{Id}$ ).

*Démonstration.* Soit un cercle  $C$  de centre  $O$  et un point  $P$ . Si  $C_2$  est un cercle perpendiculaire à  $C$  passant par  $P$ , alors nous avons vu en 22.6 que  $P' = i_C(P)$  est l'autre intersection entre  $C_2$  et la droite  $(OP)$ .

Pour construire l'image de  $P'$ , il faut un cercle perpendiculaire à  $C$  passant par  $P'$ . Le cercle  $C_2$  déjà utilisé fait l'affaire. Ensuite, la droite  $(OP')$  est la même que la droite  $(OP)$ . Donc l'image de  $P'$  est  $P$ .  $\square$

Soit  $C_1$  le cercle dans  $\mathbb{C}$  de centre 0 et de rayon 1. Nous notons  $\alpha: \mathbb{C} \rightarrow \mathbb{C}$  la dilatation de rapport  $R$ .

PROPooYHQVooVFamhr

## Proposition 22.9 ([1]).

Soit  $C$  le cercle de rayon  $R$  centré en 0 ( $C = \alpha(C_1)$ ). Alors

$$i_{\alpha(C_1)} \circ \alpha = \alpha \circ i_{C_1}. \quad (22.8)$$

*Démonstration.* Soit  $z = re^{i\theta}$ ; nous devons prouver que

$$i_C(\alpha(z)) = \alpha(i_{C_1}(z)). \quad (22.9)$$

Nous avons :

$$i_C(\alpha(z)) = i_C(Rre^{i\theta}) = R^2 \frac{1}{Rr} e^{i\theta} = R \frac{1}{r} e^{i\theta} = \alpha\left(\frac{1}{r} e^{i\theta}\right) = \alpha(i_{C_1}(z)). \quad (22.10)$$

$\square$

La définition 23.70 pour l'inversion sur  $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$  sera basée sur la proposition suivante.



**Proposition 22.10** ([1]).

Soit le cercle de rayon  $R$  et de centre  $a \in \mathbb{C}$ . Alors

$$i_C(z) = \frac{R^2}{\bar{z} - \bar{a}} + a. \quad (22.11)$$

De plus si  $t_a$  est la translation de vecteur  $a$ , nous avons la décomposition

$$i_{t_a(C_R)} = t_a \circ i_{C_R} \circ t_{-a} \quad (22.12)$$

où  $C_R$  est le cercle de rayon  $R$  centré en 0.

*Démonstration.* Si  $z \in \mathbb{C}$  et  $z'$  est son image par  $i_C$ , alors non seulement

$$\|a - z'\| \|a - z\| = R^2, \quad (22.13)$$

mais en plus  $a - z' = \lambda(a - z)$  pour un certain  $\lambda > 0$ . Cela est l'expression du fait que  $z'$  est sur la demi-droite qui joint  $a$  à  $z$ . Nous avons donc

$$\lambda \|a - z\|^2 = R^2 \quad (22.14)$$

et alors

$$\lambda = \frac{R^2}{(a - z)(\bar{a} - \bar{z})}. \quad (22.15)$$

En récrivant  $a - z' = \alpha(a - z)$  avec cette valeur de  $\lambda$  nous trouvons

$$a - z' = \frac{R^2}{a - z}, \quad (22.16)$$

ce qu'il fallait démontrer.

La décomposition demandée est une simple vérification en utilisant  $i_{C_R}(z) = \frac{R^2}{\bar{z}}$  qui découle de la proposition 22.9.  $\square$

Avant d'aller plus loin, donnons l'équation d'un cercle dans  $\mathbb{C}$ . Si  $C$  est un cercle de centre  $\omega$  et de rayon  $r$ , alors  $z \in C$  si et seulement si  $d(z, \omega) = r$ . En développant, et en passant au carré sans perte d'information (les deux membres sont positifs),  $z \in C$  si et seulement si

$$(z - \omega)(\bar{z} - \bar{\omega}) = r^2. \quad (22.17)$$

En développant nous pouvons aussi écrire l'équation du cercle sous la forme

$$z\bar{z} - z\omega - \bar{z}\omega = r^2 - \omega\bar{\omega}. \quad (22.18)$$

**Proposition 22.11** ([605]).

Soit un cercle  $C$  de centre  $O$ . L'inversion

$$i_C: \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2 \setminus \{0\} \quad (22.19)$$

transforme

- (1) les droites passant par  $O$  en elles-mêmes ;
- (2) les cercles passant par  $O$  en des droites ne passant pas par  $O$  ;
- (3) les droites ne passant pas par  $O$  en des cercles passant par  $O$  ;
- (4) les cercles ne passant pas par  $O$  en des cercles ne passant pas par  $O$ .

*Démonstration.* Point par point.

- (1) Une droite passant par  $O$  est une droite perpendiculaire à  $C$ . Par le point (3) de la définition 22.5, elle est invariante.

(2) Nous construisons successivement :

- Un cercle  $C_1$  de centre  $O_1$  et passant par  $O$ . Le but est de déterminer l'image de ce cercle.
- Le point  $P$  de  $C_1$  tel que  $[OP]$  en soit un diamètre.
- Le point  $P' = i_C(P)$ .
- La droite  $\ell$  perpendiculaire à  $(OP)$  et passant par  $P'$ .

Nous montrons maintenant que  $i_C(C_1 \setminus \{O\}) = \ell$ . Soit  $Q \in C_1 \setminus \{P, O\}$ . Nous posons

$$Q' = (QO) \cap \ell. \quad (22.20)$$

Vu que  $[OP]$  est un diamètre de  $C_1$  et que  $Q \in C_1$ , le triangle  $OPQ$  est rectangle en  $Q$ . Et étant donné que  $Q'$  est sur  $\ell$  nous savons que  $OP'Q'$  est rectangle en  $P'$ .

De plus les angles en  $O$  de ces deux triangles sont identiques (parce que c'est l'angle formé par les droites  $(OQ)$  et  $(OP)$ ) ; les triangles  $OPQ$  et  $QP'Q'$  sont donc semblables et nous pouvons utiliser le théorème de Thalès<sup>1</sup> 12.169 :

$$\frac{OP}{OQ} = \frac{OQ'}{OP'}. \quad (22.21)$$

Donc

$$\|OP\| \|OP'\| = \|OQ\| \|OQ'\|, \quad (22.22)$$

mais  $P'$  est l'image de  $P$  par l'inversion du cercle  $C$ , c'est-à-dire  $\|OP\| \|OP'\| = R^2$ . Nous en déduisons que

$$\|OQ\| \|OQ'\| = R^2, \quad (22.23)$$

c'est-à-dire que  $Q'$  est l'image de  $Q$  par  $i_C$ , et donc que

$$i_C(C_1 \setminus \{O\}) \subset \ell. \quad (22.24)$$

Pour avoir l'inclusion inverse, il faut remarquer que  $\ell$  est parallèle à la tangente à  $C_1$  en  $O$ . Donc si  $Q \in \ell$ , la droite  $(OQ)$  intersecte le cercle  $C_1$  en un point  $Q'$ . En refaisant le cheminement du résultat (22.24) à l'envers, il est loisible de prouver que  $i_C(Q') = Q$  et donc que  $\ell$  est bien incluse à l'image de  $C_1$  par  $i_C$ .

(3) Nous commençons par prouver que toutes les droites ne passant pas par  $O$  sont des images de cercles passant par  $O$ .

Nous considérons :

- une droite  $\ell$  ne passant pas par  $O$ .
- la droite  $d$ , perpendiculaire à  $\ell$  passant par  $O$
- le point  $P' = \ell \cap d$ ,
- le point  $P = i_C(P')$ ,
- le cercle  $C_1$  dont  $[OP]$  est un diamètre.

Par tout ce que nous avons fait jusqu'à présent, la droite  $\ell$  est l'image du cercle  $C_1$ . Or si  $\|OP\| = r$  alors

$$\|OP'\| = \frac{R^2}{r}. \quad (22.25)$$

Donc quelle que soit la valeur de  $\|OP'\|$  dans  $]0, \infty[$ , il existera un point  $P$  tel que le cercle passant par  $O$  et  $P$  ait pour image la droite perpendiculaire à  $(OP)$  passant par  $i_C(P)$ .

Étant donné que  $i_C$  est une involution surjective des cercles passant par  $O$  vers les droites ne passant pas par  $O$ , elle transforme également toutes les droites ne passant pas par  $O$  en un cercle passant par  $O$ .

1. Faites bien le dessin : ce n'est pas une situation de Thalès ultra-standard de collège.

(4) Pour cette partie, nous allons utiliser un peu de géométrie analytique dans  $\mathbb{C}$ <sup>2</sup>.

(4a) **Le cas centré** Nous supposons que  $C$  est centré en 0 et de rayon, 1, ce telle sorte que  $i_C(z) = 1/\bar{z}$  (proposition 22.10). Soit  $C_1$  un cercle de centre  $\omega$  et de rayon  $r$ , ne passant pas par 0, en particulier tel que  $|\omega| \neq r$ .

Si  $z \in i_C(C_1)$  alors  $i_C(z) \in C_1$  et nous avons

$$(i_C(z) - \omega) \overline{(i_C(z) - \omega)} = r^2. \quad (22.26)$$

En développant et en multipliant par  $z\bar{z}$  nous trouvons

$$\bar{\omega}z + \omega\bar{z} + z\bar{z}(r^2 - \omega\bar{\omega}) = 1. \quad (22.27)$$

ou, en posant  $s = r^2 - \omega\bar{\omega}$ ,

$$z\bar{\omega} + \bar{z}\omega + sz\bar{z} = 1. \quad (22.28)$$

En comparant avec l'équation (22.18) d'un cercle, nous voyons qu'il y a de l'espoir de montrer que  $z$  est sur un cercle centré en  $a = -\omega/s$ . En effet, en divisant par  $s$  et en substituant  $a$  comme il faut, nous trouvons

$$-z\bar{a} - \bar{z}a + a\bar{z} = \frac{1}{s}. \quad (22.29)$$

En ce qui concerne le rayon  $R$ , il doit vérifier

$$\frac{1}{s} = R^2 - a\bar{a}, \quad (22.30)$$

ou encore

$$R = \frac{1}{s} + \frac{\omega\bar{\omega}}{s^2}. \quad (22.31)$$

Nous devons vérifier que  $R > 0$ . En multipliant par  $s^2$ , il s'agit de vérifier que  $s + \omega\bar{\omega} \geq 0$ , ce qui est correct parce que  $s + \omega\bar{\omega} = r^2$ .

En résumé, si  $z \in i(C_1)$  alors  $z$  est dans le cercle  $C_2$  de centre  $-\frac{\omega}{s}$  et de rayon  $R$ . Nous voyons que ce cercle ne passe pas par 0 par exemple directement sur l'équation (22.27) qui, pour  $z = 0$  donne  $0 = 1$ .

Nous avons prouvé que  $i(C_1) \subset C_2$ . Pour prouver l'inclusion inverse, vu que  $i$  est une involution, il faut prouver  $i(C_2) \subset C_1$ . Pour cela nous écrivons l'équation qui donne  $i(z) \in C_2$  et en développant nous devons conclure que  $z \in C_1$ . Nous ne le faisons pas ici.

(4b) **Le cas de rayon non unité**

Si  $C$  est un cercle quelconque, nous écrivons l'inversion du cercle  $C$  via la formule (22.11). Si  $z \in i_C(C_1)$  alors  $i(z) \in C_1$  et nous pouvons écrire

$$(i(z) - a) \overline{(i(z) - a)} = R^2. \quad (22.32)$$

De là il faut déduire que  $z$  est sur un cercle ne passant pas par 0. Bons calculs...

□

2. Principalement parce que je ne comprends pas le raisonnement fait dans [605].

# Chapitre 23

## Espaces projectifs

Sur les espaces projectifs : [673].

### Définition 23.1.

Soit  $E$  un espace vectoriel de dimension finie sur le corps commutatif  $\mathbb{K}$ . Nous définissons sur  $E \setminus \{0\}$  la relation d'équivalence  $u \sim v$  si et seulement si  $u = \lambda v$  pour un certain  $\lambda \in \mathbb{K}$ . Cette relation est la relation de **colinéarité**. L'ensemble des classes d'équivalence de  $\sim$  est l'**espace projectif** de  $E$  et sera noté  $P(E)$ .

DEFooTPPMooTDxNpg

### Définition 23.2.

Si  $\dim E = 2$ , l'ensemble  $P(E)$  est la **droite projective**, et si  $\dim E = 3$  nous parlons du **plan projectif**.

Étant donné que tous les  $\mathbb{K}$ -espaces vectoriels de dimensions  $n + 1$  sont isomorphes à  $\mathbb{K}^{n+1}$ , nous noterons  $P_n(\mathbb{K})$  ou  $P_n$  l'espace projectif  $P(\mathbb{K}^{n+1})$ . | PgNotimesjNtMoW

### Exemple 23.3.

Si  $n = 1$  et  $\mathbb{K} = \mathbb{R}$ , l'espace projectif est l'ensemble des droites vectorielles dans le plan usuel. Il y en a une pour chaque point du type  $(x, 1)$  avec  $x \in \mathbb{R}$  et ensuite une horizontale, passant par le point  $(1, 0)$ . Nous avons donc

$$P_1(\mathbb{R}) = \{(1, 0)\} \cup \{(x, 1) \text{ tel que } x \in \mathbb{R}\}. \quad (23.1)$$

Le point  $(1, 0)$  est dit « point à l'infini ».

△

## 23.1 Sous espaces projectifs

Un **sous-espace projectif** de  $P(E)$  est une partie de la forme  $P(F)$  où  $F$  est un sous-espace vectoriel de  $E$ .

PropuqpWVx

### Proposition 23.4.

Si  $F$  et  $G$  sont des sous-espaces vectoriels de  $E$ , alors

$$P(F) \cap P(G) = P(F \cap G) \quad (23.2)$$

et nous avons

$$\dim P(F) + \dim P(G) = \dim P(F + G) + \dim P(F \cap G). \quad (23.3) \quad \text{EqNAAdWfN}$$

*Démonstration.* Nous avons

$$P(F) = \{[v] \text{ tel que } v \in F\} \quad (23.4)$$

où les crochets signifient la classe par rapport à la relation de colinéarité. Nous avons alors

$$P(F) \cap P(G) = \{[v] \text{ tel que } v \in F \cap G\} = P(F \cap G). \quad (23.5)$$

Cela prouve le premier point.

En ce qui concerne l'équation (23.3), en considérant  $\dim P(E) = \dim E - 1$  nous devons prouver l'égalité

$$\dim F + \dim G = \dim(F + G) + \dim(F \cap G) \quad (23.6)$$

concernant les dimensions des espaces vectoriels usuelles. Si nous considérons une base de  $E$  telle que  $B_1 = \{e_1, \dots, e_{k_1}\}$  est une base de  $F \cap G$ ,  $B_2 = \{e_{k_1+1}, \dots, e_{k_2}\}$  complète  $B_1$  en une base de  $F$  et  $B_3 = \{e_{k_2+1}, \dots, e_n\}$  complète  $B_1 \cup B_2$  en une base de  $G$ .

Nous avons alors

$$\dim F + \dim G = 2 \operatorname{Card}(B_1) + \operatorname{Card}(B_2) + \operatorname{Card}(B_3) \quad (23.7a)$$

$$\dim(F + G) = \operatorname{Card}(B_1) + \operatorname{Card}(B_2) + \operatorname{Card}(B_3) \quad (23.7b)$$

$$\dim(F \cap G) = \operatorname{Card}(B_1). \quad (23.7c)$$

De là la relation (23.3) se déduit immédiatement.  $\square$

**Théorème 23.5** (incidence).

Soient  $F$  et  $G$  deux sous-espaces vectoriels de  $E$  tels que

$$\dim P(F) + \dim P(G) \geq \dim P(E). \quad (23.8)$$

Alors  $P(F) \cap P(G) \neq \emptyset$ .

*Démonstration.* En utilisant les hypothèses et la proposition 23.4 nous avons

$$\dim P(E) + \dim P(G) = \dim P(F + G) + \dim P(F \cap G) \geq \dim P(E). \quad (23.9)$$

En passant aux espaces vectoriels correspondants,

$$\dim(F + G) + \dim(F \cap G) \geq \dim(E) + 1. \quad (23.10)$$

Mais nous avons aussi  $\dim(F + G) \leq \dim(E)$  et par conséquent  $\dim(F \cap G) \geq 1$ . Au final,  $\dim P(F \cap G) \geq 0$ . Cela prouve que  $P(F \cap G)$  contient au moins un élément (nous rappelons que lorsqu'un espace projectif contient un seul élément, sa dimension est zéro).  $\square$

**Exemple 23.6.**

Soient les plans  $\Pi_1 \equiv x = 0$  et  $\Pi_2 \equiv y = 0$ . Nous avons

$$P(\Pi_1) = \{[0, y, 1]\} \cup \{[0, 1, 0]\} \quad (23.11a)$$

$$P(\Pi_2) = \{[x, 0, 1]\} \cup \{[1, 0, 0]\} \quad (23.11b)$$

où le crochet signifie la classe pour la colinéarité. Ces deux droites projectives ont comme point d'intersection le point  $[0, 0, 1]$ .  $\triangle$

**Définition 23.7.**

Un **hyperplan projectif** est un sous-espace projectif de  $P(E)$  de la forme  $P(V)$  où  $V$  est un hyperplan de  $E$ .

DEFooBBMBooSVgTnn

**Définition 23.8.**

Soit  $E$  un espace vectoriel de dimension au moins 3. Nous disons que  $d \subset P(E)$  est une **droite projective** de  $P(E)$  si  $d = P(D)$  pour un plan vectoriel  $D \subset E$ .

Nous disons que trois points de  $P(E)$  sont **alignés** lorsqu'il existe une droite projective les contenant.

**23.9.**

Dans la définition 23.8 nous voyons  $P(D)$  comme inclus dans  $P(E)$  dès que  $D$  est un sous-espace vectoriel de  $E$ . Cela est possible parce que si la direction de  $v \in D$ , c'est-à-dire la classe  $[v]$  est également une direction dans  $E$ .

Le lemme suivant peut paraître idiot, mais ce qui serait sûrement idiot est de l'utiliser sans s'en rendre compte.

**Lemme 23.10.**

*Deux points dans  $P(E)$  sont toujours alignés.*

*Démonstration.* Soient deux points  $A, B \in P(E)$ . Si  $A = \pi(a)$  et  $B = \pi(b)$  alors le plan  $D$  passant par  $a, b$  et  $0$  est vectoriel et  $P(D)$  contient  $A$  et  $B$ .

Note : si  $a, b$  et  $0$  sont trois points alignés, alors  $A = B$ . Il suffit de prendre les points  $a, c$  et  $0$  où  $c \in E$  est un point quelconque non aligné avec  $0$  et  $a$ . Nous avons de toutes façons  $A = B = \pi(a)$ .  $\square$

**Lemme 23.11.**

*Trois points distincts  $A, B, C$  dans  $P(E)$  sont alignés si et seulement si il existe trois points non alignés  $a, b, c \in E$  tels que*

- (1) *le plan passant par  $a, b$  et  $c$  est vectoriel (c'est-à-dire passe par  $0$ ),*
- (2)  *$A = \pi(a), B = \pi(b)$ , et  $C = \pi(c)$ .*

*Démonstration.* Deux implications à montrer.

- (1) **Sens direct** Soient  $A, B, C$  distincts et alignés dans  $P(E)$ . Alors il existe un plan vectoriel  $D$  tel que  $A, B, C \in P(D)$ .

La condition  $A \in P(D)$  implique qu'il existe  $a \in D$  tel que  $A = \pi(a)$ . Idem pour  $B$  et  $C$ . Les points  $a, b$  et  $c$  ainsi construits sont distincts parce que  $A, B$  et  $C$  sont distincts. Si par malheur ces trois points étaient alignés, ce n'est pas grave : il suffit de remplacer  $a$  par  $\lambda a$  avec  $\lambda \neq 0$  pour qu'ils ne le soient plus (cette manipulation ne change pas le fait que le nouveau choix de point  $a$  reste dans  $D$  parce que  $D$  est vectoriel). Nous avons donc trois points non alignés  $a, b$  et  $c$  tous contenus dans  $D$ . Le plan  $D$  répond à la question.

- (2) **Sens réciproque**

Soient  $a, b$  et  $c$  non alignés dans  $E$  tels que  $A = \pi(a), B = \pi(b)$  et  $C = \pi(c)$ . Le plan  $D$  les contenant tous trois est vectoriel par hypothèse. Nous avons  $A, B, C \in P(D)$  et donc  $A, B$  et  $C$  sont alignés dans  $P(E)$ .  $\square$

**Proposition 23.12.**

*Soit  $H = P(V)$  un hyperplan projectif de  $P(E)$  et soit  $m$  hors de  $H$ . Alors toute droite projective passant par  $m$  coupe  $H$  en un et un seul point.*

*Démonstration.* Si  $\dim E = n$  nous avons  $\dim V = n - 1$ . Soit  $d = P(D)$  une droite projective passant par  $m$ , c'est-à-dire que  $D$  est de dimension 2 dans  $E$ . Si  $D \subset V$  alors  $m \in P(D) \subset P(V)$  ; or nous avons demandé que  $m$  soit hors de  $P(V)$ . Par conséquent  $D$  n'est pas inclus dans  $V$  et en particulier  $\dim(D + V) = \dim(E)$ .

Nous recopions la formule (23.3) pour notre cas :

$$\underbrace{\dim d}_{=1} + \underbrace{\dim H}_{=n-2} = \underbrace{\dim P(D + V)}_{=n-1} + \dim P(D \cap V). \quad (23.12)$$

Nous avons donc  $\dim P(D \cap V) = 0$ , ce qui signifie que l'ensemble  $P(D \cap V) = P(D) \cap P(V) = d \cap H$  contient un et un seul point.  $\square$

## 23.2 Espace projectifs comme « complétés » d'espaces affines

Soit  $E$  un espace vectoriel de dimension 2 et  $P(E)$  la droite projective correspondante, et soit  $\{e_1, e_2\}$  une base de  $E$ . Nous considérons la droite affine  $d \equiv y = 1$ . Nous avons la bijection

$$\begin{aligned} \phi: d \cup \{\infty\} &\rightarrow P(E) \\ (x, 1) &\mapsto \text{la droite vectorielle passant par } (x, 1) \\ \infty &\mapsto \text{la droite vectorielle passant par } (1, 0). \end{aligned} \tag{23.13}$$

### Lemme 23.13.

Si nous munissons l'ensemble  $d \cup \{\infty\}$  de la topologie compactifiée d'Alexandrov<sup>1</sup>, la bijection (23.13) est un homéomorphisme.

Soient maintenant les plans affines dans l'espace vectoriel  $E$  de dimension 3

$$\Pi_1 \equiv z = 0 \tag{23.14a}$$

$$\Pi_2 \equiv z = 1. \tag{23.14b}$$

Une droite (vectorielle) de  $E$  coupe  $\Pi_2$  en un et un seul point, sauf si elle est contenue dans  $\Pi_1$ . Nous avons donc une bijection

$$\begin{aligned} \phi: P(E) &\rightarrow \Pi_2 \cup P(\Pi_1) \\ d &\mapsto \begin{cases} \Pi_2 \cap d & \text{si cette intersection est non vide} \\ d & \text{sinon.} \end{cases} \end{aligned} \tag{23.15}$$

La droite projective  $P(\Pi_1)$  est la droite à l'infini du plan projectif  $P(E)$ . Nous voyons que le plan projectif  $P(E)$  peut être vu comme un plan affine ( $\Pi_2$ ) « complété » par une droite affine  $P(\Pi_1)$ . Cette dernière droite est elle-même une droite affine complétée par un point à l'infini.

Nous pouvons généraliser cette démarche en considérant un espace affine  $\mathcal{E}$  de direction  $E$  sur le corps  $\mathbb{K}$ . Nous construisons  $F = E \times \mathbb{K}$  et nous considérons un repère affine sur  $F$  tel que  $E \equiv x_{n+1} = 0$ . Nous pouvons donc identifier  $\mathcal{E}$  à l'hyperplan affine d'équation  $x_{n+1} = 1$  dans  $F$ .

Une droite vectorielle de  $F$  non contenue dans  $E$  coupe  $\mathcal{E}$  en un unique point ; nous avons donc une bijection

$$\mathcal{E} \cup P(E) \rightarrow P(F). \tag{23.16}$$

Dans ce cadre,  $P(E)$  est l'hyperplan à l'infini et nous disons que  $P(E)$  est la **complétion projective** de  $\mathcal{E}$ .

### Exemple 23.14.

Nous considérons les plans affines

$$\Pi_1 \equiv z = 0 \tag{23.17a}$$

$$\Pi_2 \equiv z = 1 \tag{23.17b}$$

et nous avons la bijection

$$P(E) = \Pi_2 \cup P(\Pi_1). \tag{23.18}$$

Un plan vectoriel  $D$  a deux possibilités : soit il coupe  $\Pi_2$  en une droite, soit il est égal à  $\Pi_1$ . Si  $D \cap \Pi_2 = d$  ( $d$  est une droite affine), alors nous avons

$$P(D) = d \cup \{\infty_D\}, \tag{23.19}$$

ce qui justifie la terminologie comme quoi  $P(D)$  est une droite dans  $P(E)$ .  $\triangle$

Soit  $E$  un espace vectoriel de dimension 3 et le plan projectif  $P(E)$ . Nous avons deux types de droites projectives :

---

1. Définition 7.185.

- (1) D'abord nous avons la droite à l'infini, donnée<sup>2</sup> par  $P(z = 0)$ .
- (2) Ensuite nous avons toutes les droites affines du plan  $z = 1$ . Chacune de ces droites est complétée par un point à l'infini.

ExempMyTmFp

**Exemple 23.15.**

Étudions un peu le second type de droites. D'abord si deux droites sont parallèles, leurs points à l'infini sont identiques. Prenons par exemple les droites  $d = \{z = 1, x = 1\}$  et  $d' = \{z = 1, x = 2\}$ . Elles décrivent les directions des vecteurs

$$\begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 2 \\ y \\ 1 \end{pmatrix}. \quad (23.20)$$

En normalisant, ce sont les vecteurs

$$\frac{1}{\sqrt{2+y^2}} \begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} \quad \text{et} \quad \frac{1}{\sqrt{5+y^2}} \begin{pmatrix} 2 \\ y \\ 1 \end{pmatrix}, \quad (23.21)$$

et toutes deux tendent vers le vecteur  $(0, 1, 0)$  pour  $y \rightarrow \infty$ .  $\triangle$

**Lemme 23.16.**

*Deux droites d'un plan projectif ont toujours une intersection.*

*Démonstration.* Si les deux droites sont des droites affines non parallèles, le résultat est évident. Si elles sont parallèles, alors l'intersection est donnée par le point à l'infini comme indiqué dans l'exemple 23.15.

Supposons que  $d$  est la droite à l'infini tandis que  $d'$  est une droite affine. Dans notre représentation usuelle du plan affine, la droite à l'infini  $d$  contient les vecteurs  $(1, y, 0)$  et le point à l'infini  $(0, 1, 0)$ . La droite affine  $d'$  a pour équation paramétriques

$$\begin{cases} x = at + c \\ y = bt + d \\ z = 1. \end{cases} \quad \begin{array}{l} (23.22a) \\ (23.22b) \\ (23.22c) \end{array}$$

Les directions données par la droite  $d'$  sont donc

$$\frac{1}{a^2t^2 + b^2t^2 + c^2 + d^2} \begin{pmatrix} at + c \\ bt + d \\ 1 \end{pmatrix} \quad (23.23)$$

Son point à l'infini est la direction du vecteur  $(a, b, 0)$ , qui est bien un point de la droite à l'infini (éventuellement son point à l'infini<sup>3</sup>).  $\square$

La plupart du temps nous considérons le plan projectif comme étant le plan affine  $z = 1$  de l'espace affine de dimension 3 complété par la droite affine  $x = 1, z = 0$ , elle-même complétée par le point  $(0, 1, 0)$ . Ce n'est évidemment pas la seule manière. Tout plan peut être considéré comme le plan à l'infini et pour une droite projective, tout point peut être considéré comme point à l'infini.

Sur la figure 23.1(a), le point à l'infini est la direction  $(1, 0)$  tandis que la direction  $(1, 1)$  n'a rien de spécial. À l'inverse sur la figure 23.1(b), la direction à l'infini est  $(1, 1)$  tandis que la direction  $(1, 0)$  est une direction usuelle.

2. Dans notre représentation usuelle du plan projectif  $z = 1$ .

3. D'accord, aller chercher le point à l'infini de la droite à l'infini, c'est chercher loin, mais n'empêche que ça existe.



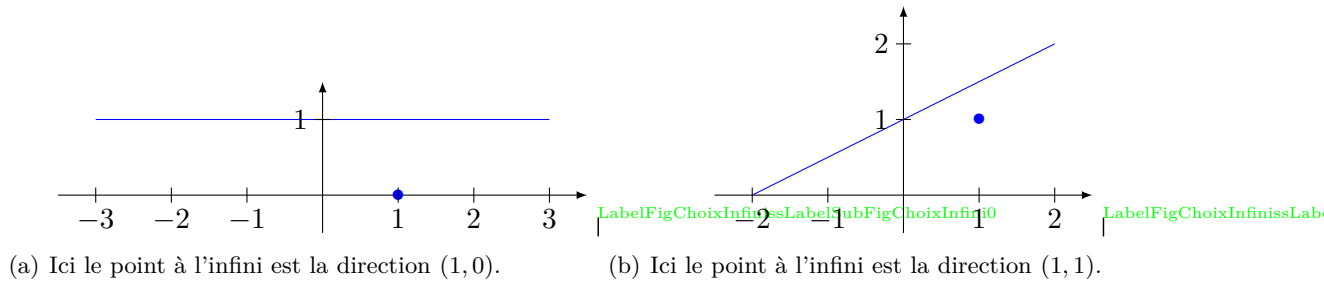


FIGURE 23.1: Deux façons de voir la droite projective. Étant donné que les points de la droite projective doivent être interprétés comme des directions (des classes d'équivalence), en réalité les deux dessins représentent les mêmes ensembles.

### Remarque 23.17.

Du point de vue de la topologie, si nous mettons celle de la compactification d'Alexandrov, tous les points de la droite projective sont équivalents.

Du point de vue de la géométrie différentielle, c'est la même chose. En effet nous pouvons mettre sur la droite projective un système de deux cartes en pensant aux angles. La première sur  $] -a, a[$  avec par exemple  $a < \pi/4$ . La seconde carte serait  $]a/2, \pi[$ . Dans ce cas la direction  $\theta = 0$  semble jouer un rôle spécial, mais il n'en est rien.

Nous pouvons également considérer les cartes  $] \pi/4 - a, \pi/4 + a[$  et  $] \pi/4 + a/2, 5\pi/4[$ . Dans ces cartes, c'est plutôt le point  $\theta = \pi/4$  qui semble différent (encore qu'il soit bien centré dans une carte).

## 23.3 Théorème de Pappus

### Théorème 23.18.

Soient deux droites  $d$  et  $d'$  dans un plan affine. Soient  $A, B, C \in d$  et  $A', B', C' \in d'$  tels que  $AB' \parallel BA'$  et  $BC' \parallel B'C$ . Alors  $AC' \parallel A'C$ .

*Démonstration.* Si  $d$  et  $d'$  ne sont pas parallèles nous considérons  $o$ , le point d'intersection. Les relations de parallélisme des hypothèses impliquent qu'il existe  $\lambda_1$  et  $\lambda_2$  tels que

$$\begin{cases} A = \lambda_1 B \\ B' = \lambda_1 A' \end{cases} \quad (23.24a)$$

$$\quad (23.24b)$$

et

$$\begin{cases} B' = \lambda_2 C' \\ C = \lambda_2 B. \end{cases} \quad (23.25a)$$

$$\quad (23.25b)$$

En substituant nous trouvons

$$\begin{cases} C = \frac{\lambda_2}{\lambda_1} A \end{cases} \quad (23.26a)$$

$$\begin{cases} A' = \frac{\lambda_2}{\lambda_1} C', \end{cases} \quad (23.26b)$$

ce qui implique que  $A'C \parallel AC'$ .

Si les droites  $d$  et  $d'$  sont parallèles, alors nous avons les translations

$$\begin{cases} B = A + x \\ A' = B' + x \end{cases} \quad (23.27a)$$

$$\quad (23.27b)$$

et

$$\begin{cases} B = C + y \\ C' = B' + y, \end{cases} \quad (23.28a)$$

$$\quad (23.28b)$$

ce qui montre que

$$\begin{cases} C = A + x - y \\ A' = C' + x - y, \end{cases} \quad \begin{matrix} (23.29a) \\ (23.29b) \end{matrix}$$

et donc que  $A'C \parallel AC'$ .  $\square$

Le théorème suivant est une version projective.

**Théorème 23.19.**

Soient  $d$  et  $d'$  deux droites projectives d'un plan projectif. Soient  $A, B, C \in d$  et  $A', B', C' \in d'$ . Alors les points  $B'C \cap C'B$ ,  $C'A \cap A'C$  et  $A'B \cap B'A$  sont alignés.

*Démonstration.* Soient  $E = B'C \cap C'B$  et  $E' = C'A \cap A'C$ . Ces deux points existent parce que deux droites projectives distinctes ont toujours un unique point d'intersection. Nous allons prendre  $EE'$  comme droite à l'infini et prouver que le point  $A'B \cap B'A$  est dessus. Étant donné que le point d'intersection de  $B'C$  et  $C'B$  est à l'infini nous avons  $B'C \parallel C'B$  (cela est un exemple de la flexibilité de la notion de parallélisme en géométrie projective). De la même façon nous avons  $C'A \parallel A'C$ .

Par le théorème de Pappus affine nous avons alors  $A'B \parallel B'A$  et par conséquent le point d'intersection est sur la droite à l'infini, c'est-à-dire sur la droite  $EE'$ .  $\square$

## 23.4 Homographies

### 23.4.1 Homographies

DEFooKWSMooXvOeEP

**Définition 23.20.**

Soient  $E$  et  $F$  deux espaces vectoriels avec leurs projections naturelles

$$\pi_E: E \setminus \{0\} \rightarrow P(E) \quad (23.30a)$$

$$\pi_F: F \setminus \{0\} \rightarrow P(F). \quad (23.30b)$$

Une application  $g: P(E) \rightarrow P(F)$  est une **homographie** si il existe un isomorphisme d'espaces vectoriels  $\bar{g}: E \rightarrow F$  tel que le diagramme

$$\begin{array}{ccc} E \setminus \{0\} & \xrightarrow{\bar{g}} & F \setminus \{0\} \\ \pi_E \downarrow & & \downarrow \pi_F \\ P(E) & \xrightarrow{g} & P(F) \end{array} \quad (23.31)$$

commute, c'est-à-dire si il existe  $\bar{g}: E \rightarrow F$  telle que

$$\pi_F(\bar{g}(v)) = g(\pi_E(v)) \quad (23.32)$$

EQooSEFWooRpiLxt

pour tout  $v \in E$ .

**Lemme 23.21.**

Si  $\bar{g}: E \rightarrow F$  est linéaire et si  $\ker \bar{g} = \{0\}$  alors l'application  $g$  définie par

$$g(\pi_E(v)) = \pi_F(\bar{g}(v)) \quad (23.33)$$

EoRIGUW

est une homographie.

*Démonstration.* Nous devons simplement vérifier que l'équation (23.33) définit bien une application. Soient  $v, w \in E$  tels que  $\pi_E v = \pi_E w$ ; nous devons montrer que

$$\pi_F \bar{g} v = \pi_F \bar{g} w. \quad (23.34)$$

EoRooIUkH

L'équation (23.34) sera vérifiée si et seulement si il existe  $\lambda \in \mathbb{R}$  tel que  $\bar{g}v = \lambda\bar{g}w$ , c'est-à-dire si et seulement si  $\bar{g}(v - \lambda w) = 0$ . Étant donné que nous supposons que le noyau de  $\bar{g}$  est réduit à  $\{0\}$ , l'équation (23.34) sera vérifiée si et seulement si  $v = \lambda w$ , ce qui signifie exactement  $\pi_E(v) = \pi_E(w)$ .  $\square$

La proposition suivante donne les premières propriétés des homographies.

PROP00GVYX00DiiBw

### Proposition 23.22.

*Quelques propriétés des homographies.*

ITEM00TION00SKjfnY

- (1) Une homographie est bijective.
- (2) Si deux espaces projectifs sont homographes, alors ils ont même dimension.
- (3) L'ensemble des homographies  $P(E) \rightarrow P(E)$  est un groupe (pour la composition).
- (4) Une homographie conserve l'alignement des points.

ITEM00IZAP00NxEigb

*Démonstration.* Nous considérons une homographie  $g: P(E) \rightarrow P(F)$ , et  $\bar{g}$  l'isomorphisme d'espaces vectoriels correspondant.

- (1) Pour l'injectivité, si  $g([v]) = g([w])$  alors en utilisant la définition d'une homographie,  $\pi_F \bar{g}v = \pi_F \bar{g}w$ , ce qui implique que  $\bar{g}v = \lambda\bar{g}w$ , et donc  $v = \lambda w$ , ce qui signifie  $[v] = [w]$ .  
Pour la surjectivité, un élément général de  $P(F)$  prend la forme  $\pi_F \bar{g}v$  pour un certain  $v \in E$ . Nous avons  $g(\pi_E v) = \pi_F \bar{g}v$ . Par conséquent l'élément  $\pi_F \bar{g}v$  est bien dans l'image de  $g$ .
- (2) Une homographie  $P(E) \rightarrow P(F)$  n'existe que si il existe un isomorphisme  $E \rightarrow F$ . Les dimensions sont donc automatiquement égales.
- (3) Il suffit de vérifier que l'application

$$\begin{aligned} \varphi: P(E) &\rightarrow P(E) \\ \pi_F \bar{g}v &\mapsto \pi_E v \end{aligned} \tag{23.35}$$

est bien définie et donne l'inverse de  $g$ .

- (4) Soient les points  $A, B, C$  alignés dans  $P(E)$ ; ils correspondent à des directions de  $E$  qui sont données par des vecteurs situés sur la même droite affine. Autrement dit, il existe trois points  $a, b, c \in E$  situés sur la même droite affine tels que  $A, B, C = \pi_E(a, b, c)$ . Les images par  $g$  sont données par  $\pi_F \bar{g}a$ ,  $\pi_F \bar{g}b$ , et  $\pi_F \bar{g}c$ .  
Étant donné qu'un isomorphisme d'espaces vectoriels conserve l'alignement affine, les points  $\bar{g}a$ ,  $\bar{g}b$  et  $\bar{g}c$  sont alignés dans  $F$ . Cela implique que les projections par  $\pi_F$  sont alignés dans  $P(F)$ .

$\square$

### 23.4.2 Le groupe projectif

DEF00WUSD00SLVKwV

#### Définition 23.23.

Le groupe des homographies de l'espace  $P(E)$  est le **groupe projectif**, noté  $\text{PGL}(E)$ .

Nous avons une surjection naturelle

$$\begin{aligned} \text{GL}(E) &\rightarrow \text{PGL}(E) \\ \bar{g} &\mapsto g \end{aligned} \tag{23.36}$$

EappNEfe

qui s'avère être un morphisme de groupes.

#### Proposition 23.24.

Nous avons l'isomorphisme de groupes

$$\frac{\text{GL}(E)}{\{\text{homothéties}\}} \simeq \text{PGL}(E). \tag{23.37}$$

*Démonstration.* Nous devons prouver que le noyau de l'application (23.36) est constitué des homothéties. Considérons un automorphisme d'espace vectoriel  $f: E \rightarrow E$  dont l'homographie associée est l'identité, et prouvons que  $f$  est une homothétie. Nous avons le diagramme commutatif suivant :

$$\begin{array}{ccc} E \setminus \{0\} & \xrightarrow{f} & E \setminus \{0\} \\ \pi_E \downarrow & & \downarrow \pi_E \\ P(E) & \xrightarrow{\text{Id}} & P(E). \end{array} \quad (23.38)$$

Pour tout vecteur  $v \in E$  nous avons  $\pi_E(v) = \pi_E(f(v))$ . Cela implique qu'il existe  $\lambda \in \mathbb{R}$  tel que  $f(v) = \lambda v$ . Tous les vecteurs de  $E$  sont donc des vecteurs propres de  $f$ . Cela n'est possible que si toutes les valeurs propres sont identiques, c'est-à-dire que  $f$  est une homothétie.  $\square$

### 23.4.3 Repères projectifs

**Définition 23.25** ([674]).

Des éléments  $\{P_i\}_{i \in I}$  sont **projectivement indépendants** si en choisissant  $v_i \in \pi^{-1}(P_i)$  nous obtenons des vecteurs  $\{v_i\}_{i \in I}$  linéairement indépendants.

DEFooPZKFooDBXtEn

**Définition 23.26.**

Soit  $E$  un espace vectoriel de dimension  $n + 1$ . Un **repère projectif** de  $P(E)$  est la donnée de  $n + 2$  points  $m_0, \dots, m_{n+1}$  tels que

- (1) les vecteurs  $m_i$ ,  $i \neq 0$ , sont les images d'une base  $\{e_i\}$  de  $E$
- (2)  $m_0 = \pi_E(e_1 + e_2 + \dots + e_{n+1})$ .

Note que si  $m_k = \pi_E(v_k)$  ( $k = 0, \dots, n + 1$ ), alors tout choix de  $n + 1$  vecteurs parmi les  $v_k$  est une base de  $E$ .

**Exemple 23.27.**

Un repère projectif de l'espace  $P(\mathbb{R}^3)$  est par exemple les éléments  $\{m_i\}_{i=1,\dots,3}$  donnés par

$$m_1 = \pi(e_1) \quad (23.39a)$$

$$m_2 = \pi(e_2) \quad (23.39b)$$

$$m_3 = \pi(e_3) \quad (23.39c)$$

$$m_0 = \pi(e_1 + e_2 + e_3). \quad (23.39d)$$

$\triangle$

**Exemple 23.28.**

Pour  $P(\mathbb{C}^2)$ , un repère projectif possible est  $m_1 = [1, 0]$ ,  $m_2 = [0, 1]$ ,  $m_0 = [1, 1]$ .  $\triangle$

**23.29.**

Pourquoi voulons nous des repères projectifs ? Pourquoi demander un quatrième élément alors que trois devraient suffire ? Le fait est que si  $E$  est de dimension 3, nous voudrions pouvoir identifier  $E$  et  $P(\mathbb{R}^3)$ .

Plus précisément, si  $E$  est de dimension  $n + 1$  et possède une base  $\{f_i\}_{i=1,\dots,n+1}$ , il existe un unique isomorphisme d'espaces vectoriels  $E \rightarrow \mathbb{R}^{n+1}$  qui envoie cette base sur la base canonique de  $\mathbb{R}^{n+1}$ . La base de  $E$  étant fixée, nous pouvons donner à un point de  $E$  les coordonnées de son image dans  $\mathbb{R}^{n+1}$  par cet isomorphisme *qui est unique*.

Dans le cas des espaces projectifs, nous voudrions avoir une unique homographie  $\phi: P(E) \rightarrow P(\mathbb{R}^{n+1})$  qui permet de donner à un point  $A \in P(E)$  les coordonnées de  $\pi^{-1}(\phi(A))$ . Bien entendu ce dernier n'est pas un élément bien défini de  $\mathbb{R}^{n+1}$  parce qu'il y a toute une droite d'éléments de  $\mathbb{R}^n$  qui se projettent sur  $\phi(A)$ .

L'idée d'imposer un point de plus est la bonne. Si nous imposons un point de plus, nous pouvons dire que les coordonnées de  $A \in P(E)$  sont celles dans  $\mathbb{R}^{n+1}$  de l'élément de  $\pi^{-1}(\phi(A))$  dont la dernière coordonnée est par exemple 1.

Nous allons maintenant mettre ça en musique.

D'abord nous donnons un exemple de non unicité.

**Exemple 23.30.**

Soit un espace vectoriel  $E$  de dimension 2 et une base  $\{b_1, b_2\}$  de  $E$ . Nous considérons également l'espace  $\mathbb{R}^2$  muni de sa base canonique  $\{e_1, e_2\}$ .

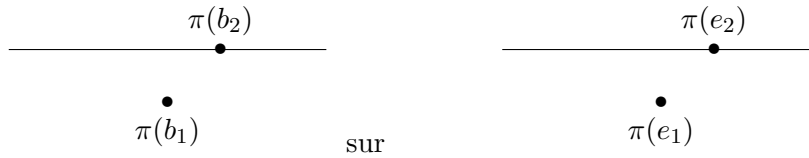
Soit une homographie  $\phi: P(E) \rightarrow P(\mathbb{R}^2)$  telle que

$$\phi(\pi(b_i)) = \pi(e_i) \quad \text{EQuoPMARoXGnKDD} \quad (23.40)$$

pour  $i = 1, 2$ . Nous allons facilement construire une autre homographie qui vérifie les mêmes conditions.

L'idée est la suivante. L'espace  $P(E)$  peut être vu comme la droite complétée  $\{(x, 1)\}_{x \in \mathbb{R}} \cup \{(1, 0)\}$  et l'espace  $P(\mathbb{R}^2)$  également. Une homographie respectant (23.40) doit envoyer le  $(1, 0)$  de  $E$  vers le  $(1, 0)$  de  $\mathbb{R}^2$  et le  $(1, 1)$  de  $E$  vers le  $(1, 1)$  de  $\mathbb{R}^2$ . Mais en ce qui concerne le reste de la droite, l'homographie peut la parcourir à la vitesse qu'elle veut.

Il faut envoyer



Soit donc une homographie  $\phi: P(E) \rightarrow P(\mathbb{R}^2)$ , et nous définissons

$$\begin{aligned} \phi': P(E) &\rightarrow P(\mathbb{R}^2) \\ \pi(xb_1 + yb_2) &\mapsto \phi(\pi(xb_1 + \lambda yb_2)) \end{aligned} \quad (23.41)$$

pour un certain  $\lambda \neq 1$ . En ce qui concerne le relèvement, l'application  $\bar{\phi}': E \rightarrow \mathbb{R}^2$  donnée par

$$\bar{\phi}'(xb_1 + yb_2) = \bar{\phi}(xb_1 + \lambda yb_2) \quad (23.42)$$

est bien définie et vérifie

$$\pi_{\mathbb{R}^2} \circ \bar{\phi}' = \phi' \circ \pi_E. \quad (23.43)$$

Donc  $\phi'$  est une homographie. De plus

$$\phi'(\pi(b_1)) = \phi(\pi(b_1)) \quad (23.44a)$$

$$\phi'(\pi(b_2)) = \phi(\pi(\lambda b_2)) = \phi(\pi(b_2)) \quad (23.44b)$$

parce que  $\pi(\lambda b_2) = \pi(b_2)$ .

Nous n'avons donc pas l'unicité. △

C'est pour rétablir cette unicité que nous demandons d'avoir un point de plus pour avoir un repère projectif. De cette façon nous aurons une unique homographie  $\phi: P(E) \rightarrow P(\mathbb{R}^{n+1})$  vérifiant  $\phi(\pi_E(b_i)) = \pi_{\mathbb{R}^{n+1}}(e_i)$  pour tout  $i = 0, \dots, n+1$ .

**Lemme 23.31.**

Soit un espace vectoriel  $E$  de dimension  $n+1$  muni de deux bases  $\{e_i\}_{i=1, \dots, n+1}$  et  $\{f_i\}_{i=1, \dots, n+1}$ . Soit un repère projectif  $\{m_0, m_i\}_{i=1, \dots, n+1}$  de  $P(E)$ .

Si  $\pi(e_i) = \pi(f_i) = m_i$  pour tout  $i = 1, \dots, n+1$  et si

$$\pi(e_1 + \dots + e_{n+1}) = \pi(f_1 + \dots + f_{n+1}) \quad (23.45)$$

alors les deux bases sont proportionnelles : il existe  $\lambda$  tel que  $f_i = \lambda e_i$  pour  $i = 1, \dots, n+1$ .

*Démonstration.* Nous avons  $\pi(e_i) = \pi(f_i)$  pour tout  $i = 1, \dots, n+1$ . Donc pour chaque  $i = 1, \dots, n+1$  il existe  $\lambda_i \in \mathbb{K}$  tel que  $e_i = \lambda f_i$ . Nous devons voir que les  $\lambda_i$  sont en réalité tous égaux.

Pour cela nous avons aussi l'égalité pour  $i = 0$  :

$$\pi(e_1 + \dots + e_{n+1}) = \pi(f_1 + \dots + f_{n+1}), \quad (23.46)$$

ce qui donne un  $\mu \in \mathbb{K}$  tel que  $e_1 + \dots + e_{n+1} = \mu(f_1 + \dots + f_{n+1})$ , c'est-à-dire

$$\lambda_1 f_1 + \dots + \lambda_{n+1} f_{n+1} = \mu f_1 + \dots + \mu f_{n+1}. \quad (23.47)$$

Du fait que les  $f_i$  forment une base, cette égalité impose à tous les  $\lambda_i$  d'être égal à  $\mu$ .  $\square$

THOooTXPVooJGigne

### **Théorème 23.32** ([675]).

Soient  $P(E)$  et  $P(F)$  deux espaces projectifs de dimensions  $n$ .

ITEMooRSIWooXbEnIT

- (1) Une homographie  $P(E) \rightarrow P(F)$  envoie un repère projectif sur un repère projectif.
- (2) Si  $(m_0, \dots, m_{n+1})$  est un repère projectif de  $P(E)$ , si  $(m'_0, \dots, m'_{n+1})$  est un repère projectif de  $P(F)$  alors il existe une unique homographie  $g: P(E) \rightarrow P(F)$  telle que  $g(m_i) = m'_i$  pour tout  $i = 0, 1, \dots, n+1$

ITEMooQXQXooDyIsxsh

*Démonstration.* Un point à la fois.

#### (1) (1)

Soit une homographie  $\phi: P(E) \rightarrow P(F)$  et un repère projectif  $\{m_0, m_1, \dots, m_{n+1}\}$  de  $P(E)$ . Nous posons  $m'_i = \phi(m_i)$  pour tout  $i = 0, \dots, n+1$ . Nous devons prouver que ces  $m'_i$  forment un repère projectif de  $P(F)$ .

D'abord pour  $i = 1, \dots, n+1$  nous avons  $m'_i = \phi(\pi_E(e_i)) = \pi_F(\bar{\phi}(e_i))$ , mais  $\{\bar{\phi}(e_i)\}_{i=1, \dots, n+1}$  est une base de  $F$  parce que  $\bar{\phi}$  est un isomorphisme d'espaces vectoriels. Donc oui : les  $m'_i$  ( $i = 1, \dots, n+1$ ) sont les projetés d'une base de  $F$ .

Nous posons au passage  $f_i = \bar{\phi}(e_i)$ . En ce qui concerne  $m_0$  nous savons que  $m_0 = \pi_E(e_1 + \dots + e_{n+1})$  et

$$m'_0 = \phi(\pi_E(e_1 + \dots + e_{n+1})) = \pi_F(\bar{\phi}(e_1 + \dots + e_{n+1})) = \pi_F(f_1 + \dots + f_{n+1}), \quad (23.48)$$

ce qui termine de montrer que  $\{m'_i\}_{i=0, \dots, n+1}$  est un repère projectif de  $P(F)$ .

#### (2) (2)

Soient un repère projectif  $(m_0, \dots, m_{n+1})$  de  $P(E)$  et un repère projectif  $(m'_0, \dots, m'_{n+1})$  de  $P(F)$ . Nous choisissons des bases  $\{e_i\}$  de  $E$  et  $\{f_i\}$  de  $F$  telles que

$$m_i = \pi_E(e_i) \quad (23.49a)$$

$$m'_i = \pi_F(f_i) \quad (23.49b)$$

pour  $i = 1, \dots, n+1$  et

$$m_0 = \pi_E(e_1 + \dots + e_{n+1}) \quad (23.50a)$$

$$m'_0 = \pi_F(f_1 + \dots + f_{n+1}). \quad (23.50b)$$

Nous considérons un isomorphisme d'espace vectoriel  $\bar{\phi}: E \rightarrow F$  tel que  $\bar{\phi}(e_i) = f_i$  pour tout  $i$ , et nous voulons définir  $\phi: P(E) \rightarrow P(F)$  par

$$\phi(\pi_E(v)) = \pi_F(\bar{\phi}(v)). \quad (23.51)$$

EQooRMYIooKcPZwD

Cela est bien défini parce que si  $\pi_E(v) = \pi_E(w)$  alors  $w = \lambda v$  et

$$\pi_F(\bar{\phi}(\lambda v)) = \pi_F(\lambda \bar{\phi}(v)) = \pi_F(\bar{\phi}(v)). \quad (23.52)$$

L'application définie par (23.51) est une homographie qui envoie  $m_i$  sur  $m'_i$  pour tout  $i = 0, \dots, n+1$ . Ceci prouve la partie « existence » du point (2).

Pour l'unicité, soient des homographies

$$\phi_1: P(E) \rightarrow P(F) \quad (23.53a)$$

$$\phi_2: P(E) \rightarrow P(F) \quad (23.53b)$$

telles que  $\phi_1(m_i) = \phi_2(m_i)$  pour tout  $i = 0, \dots, n+1$ . Soit aussi une base  $\{e_i\}_{i=1, \dots, n+1}$  de  $E$  adaptée au repère projectif, c'est-à-dire  $m_i = \pi_E(e_i)$  pour  $i = 1, \dots, n+1$  et  $\pi_E(e_1 + \dots + e_{n+1}) = m_0$ . Nous considérons aussi les isomorphismes d'espaces vectoriels  $\bar{\phi}_1$  et  $\bar{\phi}_2$ . Avec tout ce beau monde nous avons

$$\phi_1(m_i) = \pi_E(\bar{\phi}_1(e_i)) \quad (23.54a)$$

$$\phi_2(m_i) = \pi_E(\bar{\phi}_2(e_i)). \quad (23.54b)$$

Mais nous savons que  $\phi_1(m_i) = \phi_2(m_i)$ , donc nous savons que  $\pi_E(\bar{\phi}_1(e_i)) = \pi_E(\bar{\phi}_2(e_i))$ , ce qui nous fait conclure que

$$\bar{\phi}_1(e_i) = \lambda_i \bar{\phi}_2(e_i) \quad (23.55)$$

pour certaines constantes  $\lambda_i \in \mathbb{K}$ . Le même raisonnement appliqué à  $m_0$  nous donne un  $\mu \in \mathbb{K}$  tel que

$$\bar{\phi}_1(e_1) + \dots + \bar{\phi}_1(e_{n+1}) = \mu(\bar{\phi}_2(e_1) + \dots + \bar{\phi}_2(e_{n+1})). \quad (23.56)$$

En mettant l'un dans l'autre :

$$\lambda_1 \bar{\phi}_2(e_1) + \dots + \lambda_{n+1} \bar{\phi}_2(e_{n+1}) = \mu(\bar{\phi}_2(e_1) + \dots + \bar{\phi}_2(e_{n+1})). \quad (23.57)$$

Sachant que  $\{\bar{\phi}_2(e_i)\}_{i=1, \dots, n+1}$  est une base de  $F$  et nous souvenant de l'unicité de la décomposition d'un élément dans une base<sup>4</sup>, nous en déduisons que tous les  $\lambda_i$  doivent être égaux à  $\mu$ . Donc pour tout  $v \in E$  nous avons  $\bar{\phi}_1(v) = \lambda \bar{\phi}_2(v)$ .

Cela a pour conséquence que  $\phi_1 = \phi_2$ .

□

### 23.33.

Si nous avons une droite projective, trois points sont nécessaires pour créer un repère et donc pour construire une homographie de la droite sur elle-même. Soit  $E$  un espace vectoriel de dimension 2 et  $P(E)$  la droite projective qui lui est associée. Soit une homographie  $f: P(E) \rightarrow P(E)$  et  $\bar{f}: E \rightarrow E$ , l'isomorphisme d'espaces vectoriels associé (par  $f \circ \pi_E = \pi_E \circ \bar{f}$ ). Si  $\{e_1, e_2\}$  est une base de  $E$  alors l'application  $\bar{f}$  a une matrice

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{M}(2, \mathbb{K}) \quad (23.58)$$

avec  $\det A \neq 0$  parce que  $\bar{f}$  est un isomorphisme.

La plupart des points de  $P(E)$  sont représentés par des points de la forme  $(z, 1)$ . Nous voudrions savoir quelle est la direction représentée par le point  $\bar{f}(z, 1)$  ; c'est-à-dire que nous voudrions savoir  $f([z, 1])$  sous la forme  $[z', 1]$  (si possible). Nous avons

$$\bar{f}(z, 1) = (a_{11}z + a_{12}, a_{21}z + a_{22}). \quad (23.59)$$

Nous posons  $\lambda = a_{21}z + a_{22}$  et nous avons

$$\bar{f}(z, 1) = \lambda \left( \frac{a_{11}z + a_{12}}{\lambda}, 1 \right). \quad (23.60)$$

Il y a plusieurs possibilités suivant les valeurs de  $\lambda$  et de  $z$ .

- (1) Si  $\lambda = 0$  c'est que nous avons  $\bar{f}(z, 1) = (a_{11}z + a_{12}, 0)$ . L'application  $f$  envoie donc le point  $(z : 1)$  sur le point à l'infini.

---

4. Proposition 4.7.

- (2) Si  $\lambda \neq 0$ , alors  $f$  envoie le point  $(z : 1)$  vers un autre point « normal ».
- (3) Si le point de départ est le point à l'infini alors  $\bar{f}(1, 0) = (a_{11}, a_{21})$ . Cela peut être le point à l'infini ou non selon les valeurs des  $a_{ij}$ .

Dans tous les cas si nous posons

$$\begin{cases} \varphi_f(z) = \frac{a_{11}z + a_{12}}{a_{21}z + a_{22}} \\ \varphi_f(\infty) = \frac{a_{11}}{a_{21}} \end{cases} \quad (23.61a)$$

$$\quad (23.61b)$$

alors nous avons

$$\bar{f}(z, 1) = (\varphi_f(z), 1). \quad (23.62)$$

Si nous prenons la convention que  $\frac{1}{0} = \infty$  et que  $(\infty, 0)$  est le point à l'infini, alors cette application  $\varphi_f$  donne bien toutes les valeurs de  $f$ , y compris les cas à l'infini.

LEMooXNKOooBKhyzt

**Lemme 23.34** ([676]).

*Trois points distincts d'une droite projective forment un repère projectif.*

*Démonstration.* Soit une droite projective  $d = P(E)$  où  $E$  est un espace vectoriel de dimension 2 sur le corps  $\mathbb{K}$ . Soient trois points distincts  $A, B$  et  $C$  de  $d$ . Nous avons  $a, b, c \in E$  tels que  $A = \pi(a)$ ,  $B = \pi(b)$  et  $C = \pi(c)$ . Vu que  $A \neq B$ , les vecteurs  $a$  et  $b$  ne sont pas proportionnels et la partie  $\{a, b\}$  est libre dans  $E$ . Autrement dit, c'est une base<sup>5</sup>.

Il existe donc  $\alpha, \beta \in \mathbb{K}$  tels que  $c = \alpha a + \beta b$ . De plus  $\alpha$  et  $\beta$  ne sont pas nuls parce que  $C \neq A$  et  $C \neq B$ . En prenant  $a' = \alpha a$ ,  $b' = \beta b$  et  $c' = c$  nous avons :  $A = \pi(a')$ ,  $B = \pi(b')$ ,  $C = \pi(c')$  en même temps que  $\{a', b'\}$  est une base de  $E$  et  $c' = a' + b'$ . Donc  $A, B, C$  est un repère projectif de  $d = P(E)$ .  $\square$

CORooRFCZooGZiQBJ

**Corolaire 23.35** ([677]).

*Soient  $E$  et  $F$  des espaces vectoriels de dimension 2. Soient  $A_i$  ( $i = 1, 2, 3$ ) des points distincts sur  $P(E)$  et  $B_i$  distincts sur  $P(F)$ . Alors il existe une unique homographie  $P(E) \rightarrow P(F)$  portant  $A_i$  sur  $B_i$  pour tout  $i = 1, 2, 3$ .*

*Démonstration.* Il s'agit de mettre en conjonction le lemme 23.34 qui dit que les  $A_i$  forment un repère projectif de  $P(E)$  (idem : les  $B_i$  forment un repère projectif de  $P(F)$ ) et le théorème 23.32(2) qui dit que l'une va sur l'autre par une unique homographie.  $\square$

#### 23.4.4 Identifications $P(\mathbb{K}^2)$ vers $\mathbb{K} \cup \{\infty\}$

Pour rappel, une droite projective est l'espace projectif modelé sur un espace vectoriel de dimension deux (définition 23.8).

NORMooUQRUooOMIzJD

#### 23.36.

Nous allons faire un usage assez intense de bijections entre  $P(\mathbb{K}^2)$  et  $\hat{\mathbb{K}} = \mathbb{K} \cup \{\infty\}$ . Une possible est

$$\begin{aligned} \varphi_0: P(\mathbb{K}^2) &\rightarrow \mathbb{K} \cup \{\infty\} \\ [k_1, k_2] &\mapsto \begin{cases} \frac{k_1}{k_2} & \text{si } k_2 \neq 0 \\ \infty & \text{si } k_2 = 0. \end{cases} \end{aligned} \quad (23.63)$$

EQooSIJDooTHPYMb

Notons que nous utilisons ici le fait que  $\mathbb{K}$  soit commutatif, sinon il aurait fallu choisir  $k_1 k_2^{-1}$  ou  $k_2^{-1} k_1$  au lieu d'écrire gentiment  $k_1/k_2$ .

CORooFJSCooNOeAel

#### Corolaire 23.37.

*Soit une bijection  $\varphi: P(\mathbb{K}^2) \rightarrow \hat{\mathbb{K}}$ . Les points  $\{\varphi^{-1}(\infty), \varphi^{-1}(0), \varphi^{-1}(1)\}$  forment un repère projectif de  $P(\mathbb{K}^2)$ .*

5. Il convient de citer ici la proposition 4.19.



*Démonstration.* Il s'agit seulement d'une application du lemme 23.34.

Juste pour l'amusement, nous allons le prouver explicitement pour la bijection  $\varphi = \varphi_0$  donnée en (23.63). Un repère projectif est la définition 23.26. Nous avons

$$\varphi_0^{-1}(\infty) = [1, 0] = \pi_{\mathbb{K}^2}((1, 0)) \quad (23.64a)$$

$$\varphi_0^{-1}(0) = [0, 1] = \pi_{\mathbb{K}^2}((0, 1)) \quad (23.64b)$$

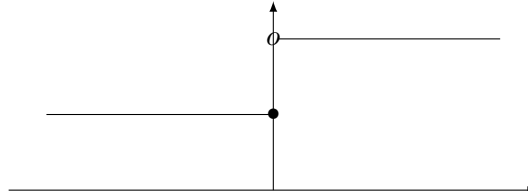
$$\varphi_0^{-1}(1) = [1, 1] = \pi_{\mathbb{K}^2}((1, 1)) \quad (23.64c)$$

Les points  $(1, 0)$  et  $(0, 1)$  forment un base de  $\mathbb{K}^2$  et nous avons bien  $(1, 1) = (1, 0) + (0, 1)$ . Donc le tout vérifie bien la définition d'un repère projectif.  $\square$

D'autre part, comme il est plus agréable de travailler avec  $\hat{\mathbb{K}}$  qu'avec  $P(\mathbb{K}^2)$  nous avons envie de voir  $\hat{\mathbb{K}}$  comme un espace projectif (qu'il n'est pas). Il y a cependant nombre d'autres identifications possibles. En voici un autre :

$$\begin{aligned} \varphi_1: P(\mathbb{K}^2) &\rightarrow \hat{\mathbb{K}} \\ [k_1, k_2] &\mapsto \begin{cases} \frac{k_2}{k_1} & \text{si } k_1 \neq 0 \\ \infty & \text{si } k_1 = 0. \end{cases} \end{aligned} \quad (23.65)$$

Vous en voulez une plus compliquée ? En voici une pour  $\mathbb{K} = \mathbb{R}$ , basée sur le dessin suivant :



La bijection associée est :

$$\begin{aligned} \varphi_d: P(\mathbb{R}^2) &\rightarrow \hat{\mathbb{R}} \\ [k_1, k_2] &\mapsto \begin{cases} \frac{k_1}{k_2} & \text{si } k_1 k_2 \leq 0 \\ 2\frac{k_1}{k_2} & \text{si } k_1 k_2 > 0 \\ \infty & \text{si } k_2 = 0. \end{cases} \end{aligned} \quad (23.66)$$

Pour mettre un peu d'ordre dans toutes ces identifications possibles, nous introduisons une classe.

### Définition 23.38.

Pour une bijection  $\varphi: P(\mathbb{K}^2) \rightarrow \hat{\mathbb{K}}$  nous définissons

$$A(\varphi) = \{ \varphi_a: P(\mathbb{K}^2) \rightarrow \hat{\mathbb{K}} \text{ tel que } \varphi_a^{-1} \circ \varphi \text{ soit une homographie} \} \quad (23.67)$$

Les classes sont assez larges parce que pour toute homographie  $\phi: P(\mathbb{K}^2) \rightarrow P(\mathbb{K}^2)$ , nous avons  $\varphi \circ \phi^{-1} \in A(\varphi)$ . Mieux, nous avons le lemme suivant.

### Lemme 23.39.

L'application

$$\begin{aligned} \psi: A(\varphi) &\rightarrow \text{PGL}(\mathbb{K}^2) \\ \varphi_a &\mapsto \varphi_a^{-1} \circ \varphi \end{aligned} \quad (23.68)$$

est une bijection.

*Démonstration.* Pour rappel,  $\text{PGL}(E)$  est le groupe des homographies de  $E$ , voir la définition 23.23.

(1) **Surjectif** Si  $\phi \in \text{PGL}(\mathbb{K}^2)$  nous avons  $\phi = \psi(\varphi \circ \phi^{-1})$ .

(2) **Injectif** Si  $\psi(\varphi_a) = \psi(\varphi_b)$  alors

$$\varphi_a^{-1} \circ \varphi = \varphi_b^{-1} \circ \varphi, \quad (23.69)$$

d'où nous déduisons  $\varphi_a^{-1} = \varphi_b^{-1}$  parce que  $\varphi$  est une bijection.

$\square$

### 23.4.5 Birapport

#### 23.40.

Tout le monde semble définir le birapport en identifiant  $P(\mathbb{K}^2)$  à  $\hat{\mathbb{K}} = \mathbb{K} \cup \{\infty\}$ . Bien entendu, personne ne semble s'être attribué la mission d'expliciter la dépendance du birapport en le choix de l'identification. Je le fais à la définition 23.42.

Mais cette définition dépend du choix d'identification  $\varphi: P(\mathbb{K}^2) \rightarrow \hat{\mathbb{K}}$ , comme le montre l'exemple 23.44. J'ai donc défini des classes d'identifications possibles  $A(\varphi)$  en 23.38. Et je démontre la proposition 23.45 que si  $\varphi_a \in A(\varphi)$  alors les birapports construits à partir de  $\varphi$  et  $\varphi_a$  sont identiques.

Question : pourquoi personne ne semble faire ce travail ? En quoi l'identification  $\varphi_0$  que tout le monde utilise est plus canonique qu'une autre ? Est-ce que l'on peut décrire simplement les classes  $A(\varphi)$  ? Le groupe qui conserve le birapport associé à  $\varphi$  est-il isomorphe au groupe qui conserve le birapport associé à  $\varphi'$  ? Quels que soient  $\varphi$  et  $\varphi'$  ?

Suis-je la seule personne au monde à m'être demandé si le birapport était un objet canonique ?

#### 23.41.

Une utilisation très intéressante du birapport dans la vidéo « The Cross-Ratio » de numberphile : [https://www.youtube.com/watch?v=ffvojZONF\\_A](https://www.youtube.com/watch?v=ffvojZONF_A)

DEFooBFSKooDwzwmO

#### Proposition-Définition 23.42.

Soit une droite projective  $d = P(E)$  et trois points distincts  $A, B$  et  $C$  sur cette droite. Soit une bijection  $\varphi: P(\mathbb{K}^2) \rightarrow \hat{\mathbb{K}}$ . Si  $X$  est un point de  $d$  alors nous nommons le **birapport** de  $X$  par rapport à  $A, B$  et  $C$  l'élément de  $\hat{\mathbb{K}}$  donné par

$$[A, B, C, X]_{\varphi} = (\varphi \circ \phi)(X) \quad (23.70)$$

où  $\phi: d \rightarrow P(\mathbb{K}^2)$  est l'unique homographie telle que

SUBEQooYWMSooyIFKQv

$$\phi(A) = \varphi^{-1}(\infty) \quad (23.71a)$$

$$\phi(B) = \varphi^{-1}(0) \quad (23.71b)$$

$$\phi(C) = \varphi^{-1}(1). \quad (23.71c)$$

*Démonstration.* Nous devons prouver qu'il existe effectivement une unique homographie vérifiant les conditions (23.71).

- (1)  $A, B, C$  est un repère projectif de  $P(E)$  Voir le lemme 23.34.
- (2)  $\varphi^{-1}(\infty), \varphi^{-1}(0), \varphi^{-1}(1)$  est un repère projectif de  $P(\mathbb{K}^2)$  Lemme 23.34 ou corolaire 23.37 au choix.
- (3) **Conclusion** Le théorème 23.32(2) nous donne existence et unicité d'une homographie  $P(E) \rightarrow P(\mathbb{K}^2)$  envoyant le premier repère sur le second.

□

#### Remarque 23.43.

La majorité des sources ne parlent pas de la dépendance du birapport en le choix de  $\varphi$  parce que tout le monde ne semble ne considérer que  $\varphi = \varphi_0$  définie en (23.63). Il est cependant naturel de se demander si la définition dépend effectivement du choix de  $\varphi$ . La réponse est oui : ça dépend du choix.

EXooYCOYooWFSfUv

**Exemple 23.44** (Une autre identification qui ne va pas bien).

Nous montrons que l'identification  $\varphi_d: P(\mathbb{K}^2) \rightarrow \hat{\mathbb{K}}$  donnée en (23.66) ne donne pas lieu au même birapport que celui de  $\varphi_0$ .

Nous travaillons le birapport sur la droite projective la plus simple :  $P(\mathbb{K}^2)$  avec  $\mathbb{K} = \mathbb{R}$ . Prenons pour la simplicité  $A = [1, 0]$ ,  $B = [0, 1]$  et  $C = [1, 1]$ . Alors l'homographie demandée dans la définition de  $[\cdot, \cdot, \cdot, \cdot]_{\varphi_0}$  est  $\phi_0 = \text{Id}$ . Par conséquent,

$$[A, B, C, [k_1, k_2]]_{\varphi_0} = (\varphi_0 \circ \phi_0)[k_1, k_2] = \varphi_0[k_1, k_2] = \frac{k_1}{k_2}. \quad (23.72)$$

En ce qui concerne le birapport défini par  $\varphi_d$  nous avons

$$\varphi_d^{-1}(\infty) = [1, 0] \quad (23.73a)$$

$$\varphi_d^{-1}(0) = [0, 1] \quad (23.73b)$$

$$\varphi_d^{-1}(1) = [1, 2], \quad (23.73c)$$

de telle sorte que nous cherchons une homographie  $\phi_d: P(\mathbb{K}^2) \rightarrow P(\mathbb{K}^2)$  telle que

$$\phi_d[1, 0] = [1, 0] \quad (23.74a)$$

$$\phi_d[0, 1] = [0, 1] \quad (23.74b)$$

$$\phi_d[1, 1] = [1, 2] \quad (23.74c)$$

L'homographie  $\phi_d[k_1, k_2] = [k_1, 2k_2]$  convient et nous avons

$$[A, B, C, D, [k_1, k_2]]_{\varphi_d} = (\varphi_d \circ \phi_d)[k_1, k_2] = \varphi_d[k_1, 2k_2] = \frac{k_1}{2k_2} \quad (23.75)$$

dès que  $k_1 k_2 < 0$ . Nous avons donc bien trouvé

$$[A, B, C, X]_{\varphi_0} \neq [A, B, C, X]_{\varphi_d}. \quad (23.76)$$

△

Le birapport n'est pas un objet tout à fait canonique parce qu'il dépend effectivement du choix de l'identification entre  $P(\mathbb{K}^2)$  et  $\hat{\mathbb{K}}$ .

PROPOoTFMQooIOQGvs

**Proposition 23.45.**

Soit une bijection  $\varphi: P(\mathbb{K}^2) \rightarrow \hat{\mathbb{K}}$ . Si  $\varphi_a \in A(\varphi)$ <sup>6</sup> alors les birapports construits sur  $\varphi$  et  $\varphi_a$  coïncident.

*Démonstration.* Soient trois points distincts  $A, B, C \in P(E)$ , et  $X \in P(E)$ . Nous avons

$$[A, B, C, X]_{\varphi} = (\varphi \circ \phi)(X) \quad (23.77)$$

où  $\phi: P(E) \rightarrow P(\mathbb{K}^2)$  est l'unique homographie telle que

$$\phi(A) = \varphi^{-1}(\infty) \quad (23.78a)$$

$$\phi(B) = \varphi^{-1}(0) \quad (23.78b)$$

$$\phi(C) = \varphi^{-1}(1). \quad (23.78c)$$

Et

$$[A, B, C, X]_{\varphi_a} = (\varphi_a \circ \phi_a)(X) \quad (23.79)$$

où  $\phi_a: P(E) \rightarrow P(\mathbb{K}^2)$  est l'unique homographie telle que

$$\phi_a(A) = \varphi_a^{-1}(\infty) \quad (23.80a)$$

$$\phi_a(B) = \varphi_a^{-1}(0) \quad (23.80b)$$

$$\phi_a(C) = \varphi_a^{-1}(1). \quad (23.80c)$$

Il est facile de voir que  $\phi_a = \varphi_a^{-1} \circ \varphi \circ \phi$ . En effet, cela est une homographie parce que  $\varphi_a \in A(\varphi)$ , et parce que la composée d'homographies est une homographie. De plus,

$$(\varphi_a^{-1} \circ \varphi \circ \phi)(A) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(\infty) = \varphi_a(\infty) \quad (23.81a)$$

$$(\varphi_a^{-1} \circ \varphi \circ \phi)(B) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(0) = \varphi_a(0) \quad (23.81b)$$

$$(\varphi_a^{-1} \circ \varphi \circ \phi)(C) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(1) = \varphi_a(1). \quad (23.81c)$$

Au final nous avons :

$$[A, B, C, X]_{\varphi_a} = (\varphi_a \circ \varphi_a^{-1} \circ \varphi \circ \phi)(X) = (\varphi \circ \phi)(X) = [A, B, C, X]_{\varphi}. \quad (23.82)$$

□

---

6.  $A(\varphi)$  définie en 23.38.

**Remarque 23.46.**

Tout le monde semble ne considérer que l'identification usuelle  $\varphi_0: P(\mathbb{K}^2) \rightarrow \hat{\mathbb{K}}$  donnée par  $\varphi_0[k_1, k_2] = k_1/k_2$ . Toute la discussion concernant la dépendance du birapport en le choix de l'identification (y compris la définition des classes  $A(\varphi)$ ) peut être sautée en disant qu'on ne considérera que  $\varphi_0$ .

Et c'est ce que nous allons faire : sauf avis contraire, nous utiliserons le birapport associé à l'identification  $\varphi_0$ .

LEM00COFT00VGKdVO

**Lemme 23.47** ([675]).

Nous avons

$$[A, B, C, X]_\varphi = \begin{cases} \infty & \text{si et seulement si } X = A \\ 0 & \text{si et seulement si } X = B \\ 1 & \text{si et seulement si } X = C. \end{cases} \quad (23.83)$$

*Démonstration.* Par définition  $[A, B, C, X]_\varphi = \varphi \circ \phi(X)$ . Nous avons donc équivalence entre les affirmations suivantes :

- $[A, B, C, X]_\varphi = \infty$
- $(\varphi \circ \phi)(X) = \infty$
- $\phi(X) = \varphi^{-1}(\infty)$
- $\phi(X) = \phi(A)$
- $A = X$

parce que  $\phi$  et  $\varphi$  sont des bijections.

Le même raisonnement tient pour les deux autres. □

PROP00KQZR00VCXPLW

**Proposition 23.48.**

*Autres petites propriétés faciles ... Soit une droite projective  $d = P(E)$  et trois points distincts  $A, B, C \in d$ .*

ITEM00OIPZ00QFFYIn

(1) *Les points  $A, B, C$  et  $X$  sont distincts si et seulement si  $[A, B, C, X] \in \mathbb{K} \setminus \{0, 1\}$ .*

ITEM00BEBE00VfiJXY

(2) *Pour tout  $k \in \hat{\mathbb{K}}$ , il existe un unique  $X \in d$  tel que  $[A, B, C, X] = k$ .*

*Démonstration.* Notons pour le point (1) que l'énoncé demande déjà que  $A, B$  et  $C$  soient distincts. Sinon le birapport n'est pas défini.

(1) (1)

Les points  $A, B$  et  $C$  sont distincts par hypothèse. Vu le lemme 23.47, pour que  $X$  soit distincts de  $A, B$  et  $C$  il faut et il suffit que le birapport ne soit ni  $\infty$  ni 1 ni 0. Donc  $\mathbb{K} \setminus \{0, 1\}$ .

(2) (2)

Nous avons  $[A, B, C, X] = \phi(X)$  où  $\phi: P(E) \rightarrow P(\mathbb{K}^2)$  est une homographie et donc une bijection par la proposition 23.22(1). Donc oui, pour tout éléments de  $P(\mathbb{K}^2)$  il existe un unique élément de  $P(E)$  dont le birapport par rapport à  $A, B$  et  $C$  soit cet élément.

Notons encore une fois que nous avons identifié  $P(\mathbb{K}^2)$  à  $\hat{\mathbb{K}}$  par la bijection (23.63). □

PROP00MGYD00HqSoJs

**Proposition 23.49** ([675]).

*Nous considérons deux droites projectives  $d$  et  $d'$  ainsi que 4 points sur chacune. Nous les nommons  $A_1, A_2, A_3, A_4 \in d$  et  $A'_1, A'_2, A'_3, A'_4 \in d'$ . Nous supposons que  $A_1, A_2, A_3$  sont distincts et que  $A'_1, A'_2, A'_3$  également. Alors il y a équivalence entre*

ITEM00IDKB00XHnNDi

(1) *Il existe une homographie  $\phi: d \rightarrow d'$  telle que  $\phi(A_i) = A'_i$  pour  $i = 1, 2, 3, 4$ ,*

ITEM00CDWA00IckJwT

(2) *égalité des birapports :*

$$[A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4]. \quad (23.84)$$

Dans ce cas, l'homographie est unique.

*Démonstration.* Nous divisons la preuve en trois parties évidentes.

(1) **(1) implique (2)**

Nous avons une homographie  $\mu': d' \rightarrow \hat{\mathbb{K}}$  telle que  $\mu'(A'_1) = \infty$ ,  $\mu'(A'_2) = 0$  et  $\mu'(A'_3) = 1$ . En composant<sup>7</sup> avec l'homographie  $\phi: d \rightarrow d'$  de l'hypothèse nous avons une homographie  $\mu' \circ \phi: d \rightarrow \hat{\mathbb{K}}$  qui vérifie

$$(\mu' \circ \phi)(A_1) = \mu'(A'_1) = \infty \quad (23.85a)$$

$$(\mu' \circ \phi)(A_2) = \mu'(A'_2) = 0 \quad (23.85b)$$

$$(\mu' \circ \phi)(A_3) = \mu'(A'_3) = 1, \quad (23.85c)$$

ce qui signifie que  $\mu' \circ \phi$  est l'homographie qui définit le birapport sur  $d$ . Par conséquent

$$[A_1, A_2, A_3, A_4] = (\mu' \circ \phi)(A_4) = \mu'(A'_4) = [A'_1, A'_2, A'_3, A'_4]. \quad (23.86)$$

(2) **(2) implique (1)**

La partie  $\{A_1, A_2, A_3\}$  est un repère projectif de  $d$  par le lemme 23.34. Idem pour  $\{A'_1, A'_2, A'_3\}$ . Nous considérons les homographies  $\mu: d \rightarrow \hat{\mathbb{K}}$  et  $\mu': d' \rightarrow \hat{\mathbb{K}}$  définissant les birapports par rapport à ces repères. Ces homographies vérifient, en utilisant l'hypothèse :

$$\mu(A_4) = [A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4] = \mu'(A'_4). \quad (23.87)$$

Et de plus

$$\mu(A_1) = [A_1, A_2, A_3, A_1] = \infty = \mu'(A'_1) \quad (23.88a)$$

$$\mu(A_2) = [A_1, A_2, A_3, A_2] = 0 = \mu'(A'_2) \quad (23.88b)$$

$$\mu(A_3) = [A_1, A_2, A_3, A_3] = 1 = \mu'(A'_3). \quad (23.88c)$$

Autrement dit :  $\mu(A_i) = \mu'(A'_i)$  pour tout  $i = 1, 2, 3, 4$ . Nous nous inspirons de ce diagramme :

$$\begin{array}{ccc} d & \xrightarrow{\phi} & d' \\ & \searrow \mu \quad \swarrow \mu' & \\ & \mathbb{K} \cup \{\infty\} & \end{array} \quad (23.89)$$

La composée  $\phi = \mu'^{-1} \circ \mu$  vérifie

$$(\mu'^{-1} \circ \mu)(A_i) = A'_i \quad (23.90)$$

pour tout  $i$ , et est une homographie.

(3) **Unicité**

Le fait qu'une homographie vérifiant  $\phi(A_i) = A'_i$  pour  $i = 1, 2, 3$  soit unique découle du fait qu'il existe une unique homographie portant un repère projectif sur un autre. A fortiori la condition  $\phi(A_4) = A'_4$  ne retire rien à l'unicité.

□

**Théorème 23.50** ([675]).

Une bijection entre deux droites projectives est une homographie si et seulement si elle conserve le birapport.

*Démonstration.* Chacun des deux sens séparément.

7. Proposition 23.22(3), la composition est encore une homographie.

(1)  $\Rightarrow$ 

Soit une homographie  $\phi: d \rightarrow d'$  entre deux droites projectives. Nous devons prouver que pour tout choix 4 points  $A, B, C, X$  dans  $d$  (dont  $A, B$  et  $C$  sont distincts) nous avons

$$[A, B, C, X] = [\phi(A), \phi(B), \phi(C), \phi(X)]. \quad (23.91)$$

Nous nommons  $\mu: d \rightarrow \hat{\mathbb{K}}$  l'homographie qui donne le birapport sur  $d$  par rapport à  $A, B$  et  $C$ , et  $\mu': d' \rightarrow \hat{\mathbb{K}}$  celle qui donne le birapport sur  $d'$  par rapport à  $\phi(A), \phi(B), \phi(C)$ . Voici un diagramme de la situation :

$$\begin{array}{ccc} d & \xrightarrow{\phi} & d' \\ & \searrow \mu \quad \swarrow \mu' & \\ & \mathbb{K} \cup \{\infty\} & \end{array} \quad (23.92)$$

Nous prouvons maintenant que  $\mu' = \mu \circ \phi^{-1}$ . En effet :

$$(\mu \circ \phi^{-1})(\phi(A)) = \mu(A) = \infty \quad (23.93a)$$

$$(\mu \circ \phi^{-1})(\phi(B)) = \mu(B) = 0 \quad (23.93b)$$

$$(\mu \circ \phi^{-1})(\phi(C)) = \mu(C) = 1. \quad (23.93c)$$

Par conséquent le birapport à droite dans (23.91) peut se calculer à l'aide de  $\mu \circ \phi^{-1}$  :

$$[\phi(A), \phi(B), \phi(C), \phi(X)] = \mu'(\phi(X)) = (\mu \circ \phi^{-1})(\phi(X)) = \mu(X) = [A, B, C, X]. \quad (23.94)$$

La première implication est prouvée.

(2)  $\Leftarrow$ 

Soit une bijection  $f: d \rightarrow d'$  conservant le birapport, ainsi que trois points distincts  $A, B$  et  $C$  dans  $d$ . Vu que  $f$  est une bijection les points  $f(A), f(B)$  et  $f(C)$  sont distincts dans  $d'$ . Par le lemme 23.34 et le théorème 23.32(2), il existe une unique homographie  $\phi: d \rightarrow d'$  telle que  $\phi(A) = f(A), \phi(B) = f(B)$  et  $\phi(C) = f(C)$ . Pour tout  $X \in d$  nous avons

$$[f(A), f(B), f(C), f(X)] = [A, B, C, X] \quad (23.95a)$$

$$= [\phi(A), \phi(B), \phi(C), \phi(X)] \quad (23.95b)$$

$$= [f(A), f(B), f(C), \phi(X)]. \quad (23.95c)$$

Justifications :

- (23.95a) parce que  $f$  conserve le birapport par hypothèse.
- (23.95b) parce que  $\phi$  conserve le birapport étant une homographie (c'est le premier sens du présent théorème)

Nous nommons  $\mu': d' \rightarrow \hat{\mathbb{K}}$  l'homographie donnant le birapport par rapport aux points  $f(A), f(B), f(C)$ . Alors le résultat (23.95) se lit

$$\mu'(f(X)) = \mu'(\phi(X)). \quad (23.96)$$

Mais comme  $\mu'$  est une bijection (proposition 23.22(1)) cela implique  $f(X) = \phi(X)$ . Vu que nous avons fait ce raisonnement pour un  $X$  quelconque dans  $d$  nous avons  $f = \phi$ , ce qui prouve que  $f$  est une homographie. □

### Lemme 23.51.

Soient  $a, b, c$  distincts sur la droite projective  $D = P(E)$ . Soient  $x, y \in E$  tels que  $\pi_E(x) = a, \pi_E(y) = b, \pi_E(x + y) = c$ . Alors

$$d = \pi_E(\lambda x + \mu y) \quad (23.97)$$

si et seulement si

$$[a, b, c, d] = \pi_{\mathbb{K}^2}(\lambda, \mu). \quad (23.98)$$

*Démonstration.* Étant donné que  $a$  et  $b$  sont distincts, les vecteurs  $x$  et  $y$  forment une base de  $E$ . Soit  $f: E \rightarrow \mathbb{K}^2$  un isomorphisme qui envoie  $(x, y)$  sur  $e_1, e_2$  où  $e_i$  sont les vecteurs de base de  $\mathbb{K}^2$ . Ensuite nous considérons  $g: P(E) \rightarrow P(\mathbb{K}^2)$ , l'homographie associée à  $f$ . Par définition  $g(\pi_E z) = \pi_{\mathbb{K}^2}(f(z))$ . Par  $f$  nous avons

$$a \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad b \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (23.99)$$

Donc par  $g$  nous avons

$$a \mapsto \infty \quad b \mapsto 0. \quad (23.100)$$

Nous avons aussi  $f(\lambda x + \mu y) = (\lambda, \mu)$  et

$$g(c) = g(\pi_E(x + y)) \quad (23.101a)$$

$$= \pi_F f(x + y) \quad (23.101b)$$

$$= \pi_F(f(x) + f(y)) \quad (23.101c)$$

$$= \pi_F \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (23.101d)$$

$$= 1. \quad (23.101e)$$

La dernière égalité est le fait que la direction  $(1, 1)$  dans  $\mathbb{R}^2$  est représentée par le point  $x = 1$  sur la droite  $y = 1$  qui est notre « représentation » de la droite affine. L'application  $g$  a donc toutes les propriétés qu'il faut pour être l'application qui définit le birapport. Nous avons donc bien  $g(d) = [a, b, c, d]$ .

D'une part si  $d = \pi_E(\lambda x + \mu y)$  alors

$$g(d) = \pi_{\mathbb{K}^2} f(\lambda x + \mu y) = \pi_{\mathbb{K}^2}(\lambda, \mu). \quad (23.102)$$

Dans l'autre sens si  $[a, b, c, d] = \pi_{\mathbb{K}^2}(\lambda, \mu)$  alors supposons que  $g(d) = \pi_{\mathbb{K}^2}(\lambda, \mu)$  avec  $d = \pi_E(v)$  alors

$$g\pi_E v = \pi_{\mathbb{K}^2} f(v), \quad (23.103)$$

ce qui implique  $f(v) = \alpha(\lambda, \mu)$  pour un certain  $\alpha \in \mathbb{K}$ . Par conséquent  $v = \alpha(\lambda x + \mu y)$  et  $d = \pi_E(\lambda x + \mu y)$ .  $\square$

## 23.5 Coordonnées homogènes

Soit  $E$  un espace vectoriel de dimension  $n + 1$  et une base  $\{e_0, \dots, e_n\}$  de  $E$ . Soit  $M \in P(E)$  et  $u \in E$  un élément engendrant  $M$ . Au point  $M$  nous voudrions associer les coordonnées  $(x_0, \dots, x_n)$  de  $u$  dans  $E$ . Notons que toutes les coordonnées de  $u$  ne sont jamais nulles en même temps parce que  $u$  doit indiquer une direction. Nous savons par ailleurs que les coordonnées  $(x_0, \dots, x_n)$  indiquent le même point de  $P(E)$  que les coordonnées  $(x'_0, \dots, x'_n)$  si et seulement si  $x_i = \lambda x'_i$ .

DEFooLWMHooMWxAFq

### Définition 23.52.

La classe d'équivalence de  $(x_0, \dots, x_n)$  est la **coordonnée homogène** de  $M$ . Nous la notons  $(x_0 : \dots : x_n)$ .

Si nous avons une base  $\{e_i\}$  de  $\mathbb{R}^n$  nous associons à  $M \in P(E)$  les coordonnées  $(X : Y : T)$ . Mais si on prend la base  $\{2e_1, e_2, \dots, e_n\}$ , les coordonnées du même point deviennent  $(X/2 : Y : T)$  alors que du point de vue de l'espace projectif, rien n'a été changé : la classe de  $e_1$  est la même que celle de  $2e_1$ . Les coordonnées homogènes<sup>8</sup> ne sont donc pas intrinsèques.

8. Définition 23.52.

### 23.5.1 Curiosité : matrice de translation

Si  $E$  est un espace vectoriel, l'espace projectif  $P(E)$  est l'ensemble des classes d'équivalence dans  $E$  pour la relation  $v \sim \lambda v$  pour tout  $\lambda \neq 0$ .

Il se fait que l'étude de  $\mathbb{R}^3$  peut être fait à partir de  $\mathbb{R}^4$  en considérant les coordonnées homogènes sur  $P(\mathbb{R}^4)$ . Plus précisément, nous considérons

$$\begin{aligned}\varphi: \mathbb{R}^3 &\rightarrow P(\mathbb{R}^4) \\ (x, y, z) &\mapsto [(x, y, z, 1)].\end{aligned}\tag{23.104}$$

Cela est injectif mais pas surjectif parce que les éléments de la forme  $[(x, y, z, 0)]$  ne sont pas atteints. Ces éléments sont alors dits « à l'infini ».

Nous aurions pu placer  $\mathbb{R}^3$  dans  $\mathbb{R}^4$  de nombreuses autres manières ; chacune aurait donné une notion différente de « point à l'infini ».

Nous allons maintenant montrer une petite curiosité qui a une grande importance en informatique, lors de la manipulation d'objets 3D. Nous considérons la bijection

$$\begin{aligned}\varphi: \mathbb{R}^3 &\rightarrow \mathbb{R}^4 \\ (x, y, z) &\mapsto (x, y, z, 1).\end{aligned}\tag{23.105}$$

Soit l'opérateur de translation  $T_a: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ . En considérant la matrice

$$T_a^h = \begin{pmatrix} 1 & 0 & 0 & a_x \\ 0 & 1 & 0 & a_y \\ 0 & 0 & 1 & a_z \\ 0 & 0 & 0 & 1 \end{pmatrix},\tag{23.106}$$

nous avons

$$T_a = \varphi^{-1} \circ T_a^h \circ \varphi.\tag{23.107}$$

Autrement dit, ce passage de  $\mathbb{R}^3$  à  $\mathbb{R}^4$  permet de voir les translations comme des matrices, et c'est bien pratique.

Si  $R: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  est une rotation, la matrice correspondante sur  $\mathbb{R}^4$  est

$$R^h = \begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix}.\tag{23.108}$$

Elle se combine assez bien avec une translation parce que le produit donne

$$T_a^h R^h = \begin{pmatrix} R & \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \\ (0 & 0 & 0) & 1 \end{pmatrix}.\tag{23.109}$$

C'est-à-dire que la translation et la rotation restent assez visible dans la matrice composée.

Note : pour la composition  $R^h T_a^h$ , c'est beaucoup moins vrai.

### 23.5.2 Dualité

Soit  $E$  un espace vectoriel de dimension  $n + 1$ . Une forme linéaire non nulle est un élément de  $E^*$ , mais aussi un représentant d'un élément de  $P(E^*)$ .

Le noyau d'une forme linéaire  $\omega$  est un hyperplan. Le noyau de la forme linéaire  $\lambda\omega$  étant le même hyperplan, l'hyperplan est donné par toute la classe de  $\omega$  dans  $P(E^*)$ . Nous avons donc une bijection

$$P(E^*) \leftrightarrow \{\text{hyperplans vectoriels de } E\}.\tag{23.110}$$

Soit un espace vectoriel  $E$  de dimension 3 muni d'une base  $\{e_1, e_2, e_3\}$  à partir de laquelle nous construisons la base duale  $\{e_1^*, e_2^*, e_3^*\}$  de l'espace dual  $E^*$ . À un élément  $m \in P(E^*)$  nous associons la droite

$$H_m = \{(X : Y : T) \text{ tel que } m(X, Y, T) = 0\}\tag{23.111}$$



dans  $P(E)$ . Si les coordonnées homogènes de  $m$  étaient  $(u : v : w)$  alors l'équation de la droite  $H_m$  est

$$uX + vY + wT = 0. \quad (23.112)$$

En effet si  $\omega \in E^*$  est un représentant de  $m$  alors  $\omega = \lambda(ue_1^* + ve_2^* + we_3^*)$  et l'équation (23.112) est indépendante de  $\lambda$  ainsi que du choix du représentant dans  $E$  du point  $(X : Y : T)$  dans  $P(E)$ .

Si les points  $m_1$  et  $m_2$  sont distincts dans  $P(E^*)$ , ils donnent deux droites  $m_1(X, Y, T) = 0$  et  $m_2(X, Y, T) = 0$ . Les points de la droite qui joint  $m_1$  à  $m_2$  dans  $P(E^*)$  sont de la forme  $\lambda m_1 + \mu m_2$  et ils sont associés à l'équation

$$\lambda m_1(X, Y, T) + \mu m_2(X, Y, T) = 0 \quad (23.113)$$

qui sont encore des droites dans  $P(E)$ . Toutes ces droites passent par le point d'intersection des droites associées à  $m_1$  et  $m_2$ . Nous avons donc

$$\bigcap_{\lambda, \mu} H_{\lambda m_1 + \mu m_2} = H_{m_1} \cap H_{m_2}. \quad (23.114)$$

### Lemme 23.53.

*L'application*

$$\begin{aligned} P(E^*) &\rightarrow \{\text{droites dans } P(E)\} \\ m &\mapsto H_m \end{aligned} \quad (23.115)$$

*est une bijection.*

*Démonstration.* Une droite dans  $P(E)$  est donnée en coordonnées homogènes par une équation  $aX + bY + cT = 0$ . Cette droite est décrite par le point  $(a : b : c)$  dans  $P(E^*)$ . Ce dernier correspond à la direction de la forme  $ae_1^* + be_2^* + ce_3^*$ . Cela prouve que l'application est surjective.

Pour l'injectivité, si  $m_1 \neq m_2$  dans  $P(E^*)$ , les formes  $\omega_1$  et  $\omega_2$  associées dans  $E^*$  ne sont pas multiples l'une de l'autre. Donc les équations

$$a_1X + b_1Y + z_1T = 0 \quad (23.116)$$

et

$$a_2X + b_2Y + z_2T = 0 \quad (23.117)$$

n'ont pas de solutions communes et décrivent donc des droites distinctes.  $\square$

LemjXywjH

### Lemme 23.54.

*Trois points distincts  $m_1$ ,  $m_2$  et  $m_3$  dans  $P(E^*)$  sont alignés si et seulement si les droites  $H_{m_1}$ ,  $H_{m_2}$  et  $H_{m_3}$  sont distinctes et concourantes.*

*Démonstration.* Supposons avoir trois points alignés, c'est-à-dire

$$m_3 = m_1 + \mu(m_2 - m_1). \quad (23.118)$$

Soit  $X : Y : T$  le point d'intersection de  $H_{m_1}$  avec  $H_{m_2}$ . Alors  $m_1(X, Y, T) = m_2(X, Y, T) = 0$ . En tenant compte de (23.118) nous avons alors évidemment  $m_3(X, Y, T) = 0$ .

Supposons maintenant que les trois droites  $H_{m_i}$  soient concourantes. Nous avons donc un point  $(X : Y : T)$  dans  $P(E)$  tel que  $m_i(X, Y, T) = 0$ . Si  $m_i$  est la classe de  $a_ie_1^* + b_ie_2^* + c_ie_3^*$  alors nous avons le système

$$\begin{cases} a_1X + b_1Y + c_1T = 0 & (23.119a) \\ a_2X + b_2Y + c_2T = 0 & (23.119b) \\ a_3X + b_3Y + c_3T = 0. & (23.119c) \end{cases}$$

Afin que cela ait une solution non triviale nous devons avoir

$$\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = 0, \quad (23.120)$$

c'est-à-dire que les points  $(a_i, b_i, c_i)$  soient alignés.  $\square$

En tenant compte de ce qui a été dit, une droite dans  $P(E^*)$  est constituée de points qui fournissent des droites concourantes dans  $P(E)$ . Donc une droite de  $P(E^*)$  se caractérise par un point de  $P(E)$  (l'intersection) de la façon suivante. Un point  $M_d \in P(E)$  donne lieu à un **faisceau de droites** passant par  $M_d$ . Chacune de ces droites donne lieu à un point de  $P(E^*)$  et tous ces points sont alignés. Nous avons ainsi construit la droite  $d$  dans  $P(E^*)$  correspondante au point  $M_d$  de  $P(E)$ .

### 23.5.3 Polynômes

Soit l'espace projectif de dimension  $n$  avec ses coordonnées homogènes  $(X_0 : \dots : X_n)$ . Nous considérons l'espace affine  $H \equiv X_n = 1$  dans l'espace vectoriel  $E$  de dimension  $n + 1$ . Nous considérons pour  $H$  un repère affine ayant pour origine le point  $(0, \dots, 0, 1)$ . Considérons un polynôme homogène  $P$  sur le corps  $\mathbb{K}$ . L'équation

$$P(X_0, \dots, X_n) = 0 \quad (23.121)$$

sur l'espace vectoriel  $E$  descend immédiatement à l'espace projectif : étant donné que  $P$  est homogène nous avons  $P(u) = 0$  si et seulement si  $P(\lambda u) = 0$ .

Nous essayons de décrire l'ensemble  $A$  des points de  $P(E)$  satisfaisant  $P(X_0, \dots, X_n) = 0$ . Nous savons que les éléments de  $P(E)$  ont chacun un représentant soit dans  $H$  soit sur la droite à l'infini. Ceux de  $A$  ayant un représentant dans  $H$  sont d'équation

$$Q(x_0, \dots, x_{n-1}) = 0 \quad (23.122)$$

où  $Q$  est le polynôme donné par  $Q(x_0, \dots, x_{n-1}) = P(x_0, \dots, x_{n-1}, 1)$ . Les points de  $A$  ayant un représentant sur la droite à l'infini s'obtiennent par l'équation

$$R(x_0, \dots, x_{n-1}) = 0 \quad (23.123)$$

où  $R$  est le polynôme donné par  $R(x_0, \dots, x_{n-1}) = P(x_0, \dots, x_{n-1}, 0)$ .

#### Exemple 23.55.

Nous considérons la conique projective

$$X^2 - XT - Y^2 - T^2 = 0. \quad \text{EqLeQIN} \quad (23.124)$$

Elle est décomposée en deux parties : une dans l'espace affine « normale » et une à l'infini. La première s'obtient en posant  $T = 1$  dans (23.124) :

$$x^2 - x - y^2 - 1 = 0. \quad \text{EqGHGJ} \quad (23.125)$$

L'autre est obtenue en posant  $T = 0$  :

$$x^2 - y^2 = 0. \quad (23.126)$$

La partie à l'infini est donc composée de deux points :  $(1 : 1 : 0)$  et  $(1 : -1 : 0)$ .

Le graphique de l'équation (23.125) est donné à la figure 23.2. Nous y voyons que les asymptotes sont effectivement données par les directions  $(1, 1)$  et  $(1, -1)$  dans le plan.

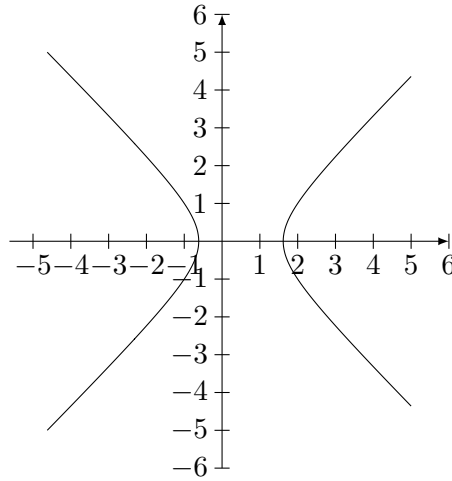
△

Nous pouvons tenter de faire l'exercice inverse : considérer une conique dans  $\mathbb{R}^2$ , la voir comme une partie d'une conique dans l'espace projectif et trouver les points à l'infini qui la complètent.

#### Exemple 23.56.

La droite projective usuelle est donnée par la droite affine  $y - 1 = 0$ . L'homogénéisation donne  $y - z = 0$  et par conséquent la partie à l'infini est donnée par  $y = 0$ , c'est-à-dire la direction  $(1, 0)$  comme il se doit.

△

FIGURE 23.2: Le graphique de  $x^2 - x - y^2 - 1 = 0$ .

LabelFigProjPoly

**Exemple 23.57.**

Prenons la conique

$$x^2 + xy + y^3 - 2 = 0. \quad (23.127)$$

D'abord nous homogénéisons cette équation pour la voir dans  $\mathbb{R}^3$  :

$$x^2z + xyz + y^3 - 2z^3 = 0. \quad (23.128)$$

Les points à l'infini sont ceux qui correspondent à  $z = 0$ , c'est-à-dire la droite donnée en coordonnées homogènes par  $(1 : 0 : 0)$ .  $\triangle$

**23.6 La sphère de Riemann  $P_1(\mathbb{C})$** 

DEFooSZGNooTzFYbh

**Définition 23.58.**

La **sphère de Riemann** est l'espace projectif modélisé sur  $\mathbb{C}^2$  : en vertu des notations données à la page 2067, c'est

$$P_1(\mathbb{C}) = P(\mathbb{C}^2). \quad (23.129)$$

L'ensemble  $P_1(\mathbb{C})$  est le quotient  $\mathbb{C}^2 \setminus \{(0,0)\} / \sim$  où  $\sim$  est la relation d'équivalence de  $\mathbb{C}$ -colinéarité dans  $\mathbb{C}^2$ .

LEMooKWZDooEIraSJ

**Lemme 23.59.**

L'application

$$\begin{aligned} \varphi_0 : P_1(\mathbb{C}) &\rightarrow \mathbb{C} \cup \{\infty\} \\ [z_1, z_2] &\mapsto \begin{cases} \frac{z_1}{z_2} & \text{si } z_2 \neq 0 \\ \infty & \text{si } z_2 = 0 \end{cases} \end{aligned} \quad (23.130)$$

EQooKJIZooZihzuU

est une bijection qui respecte la conjugaison complexe :  $\varphi_0([z_1, z_2]^*) = \varphi_0([z_1, z_2])^*$ .

*Démonstration.* Notons d'abord que la définition a un sens parce que si un représentant de  $[z_1, z_2]$  est de la forme  $(z, 0)$  alors ils sont tous de cette forme. L'affirmation «  $z_1 \neq 0$  dans  $[z_1, z_2]$  » a donc un sens.

(1) **Injectif** Supposons  $\varphi_0([z_1, z_2]) = \varphi_0([t_1, t_2])$ .

Si les deux membres sont égaux à  $\infty$  alors nous avons  $z_2 = t_2 = 0$ , et alors avec  $\lambda = z_1/t_1$  nous avons  $(z_1, z_2) = \lambda(t_1, t_2)$ , ce qui prouve que  $[z_1, z_2] = [t_1, t_2]$ .

Si les deux membres sont égaux à zéro alors  $z_1 = t_1 = 0$  et le même raisonnement tient.

Sinon nous avons  $z_1/z_2 = t_1/t_2$  où tous les nombres sont non nuls. Cela donne

$$z_2 = \frac{t_2}{t_1} z_1, \quad (23.131)$$

et donc

$$\frac{t_1}{z_1}(z_1, z_2) = (t_1, t_2), \quad (23.132)$$

qui montre qu'au niveau des classes,  $[z_1, z_2] = [t_1, t_2]$ .

## (2) Surjectif

Nous avons

$$\infty = \varphi_0([1, 0]) \quad (23.133)$$

et si  $z \neq \infty$  nous avons  $z = \varphi_0([z, 1])$ .

□

### 23.6.1 Éléments de géométrie dans $P_1(\mathbb{C})$

Étant donné que nous sommes partis pour faire de la géométrie dans  $\mathbb{C}$  et même dans  $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ , autant nous armer des équations de cercles et de droites dans  $\mathbb{C}$ , ainsi que de quelques notions adjacentes.

#### Remarque 23.60.

La définition 23.2 parle de plan et de droites projectives. Ici nous ne sommes pas dans ce cadre parce que nous travaillons sur  $P_1(\mathbb{C})$  où  $\mathbb{C}$  n'est certainement pas un espace de dimension 3. Les droites dont nous allons parler ne sont pas des droites projectives avec leur point à l'infini.

#### 23.6.1.1 Équation complexe d'une droite

L'équation d'une droite dans  $\mathbb{R}^2$  est  $d \equiv ax + by = c$  avec  $a, b, c \in \mathbb{R}$  et  $a, b$  non nuls en même temps. En posant  $z = x + iy$  nous voulons exprimer l'équation en termes de  $z$  au lieu de  $x$  et  $y$ . Nous avons[678]

$$x = \frac{z + \bar{z}}{2}, \quad y = \frac{z - \bar{z}}{2i}, \quad (23.134)$$

et nous pouvons écrire  $d \equiv a(z + \bar{z}) - ib(z - \bar{z}) = 2c$ , ou encore  $d \equiv (a - bi)z + (a + ib)\bar{z} = 2c$ . En posant  $\omega = a + bi \in \mathbb{C}^*$  et  $k = 2c \in \mathbb{R}$  nous avons l'équation

$$\bar{\omega}z + \omega\bar{z} = k. \quad (23.135)$$

#### Définition 23.61.

Une **droite** est une partie de  $\hat{\mathbb{C}}$  de la forme

$$d(\omega, k) = \{z \in \mathbb{C} \text{ tel que } \bar{\omega}z + \omega\bar{z} = k\} \cup \{\infty\} \quad (23.136)$$

avec  $\omega \in \mathbb{C}^*$  et  $k \in \mathbb{R}$ .

Dans  $\hat{\mathbb{C}}$ , toutes les droites contiennent le point  $\infty$ .

#### ⚠ Avertissement/question au lecteur !! 23.62

La proposition 22.11 montre que toute inversion transforme un cercle-droite en un cercle-droite, nonobstant d'accepter de prolonger toute droite par  $\infty$ .

Est-ce que l'on peut dire que toutes les droites contiennent le point  $\infty$  ?

En donnant  $\infty$  à toutes les droites et à aucun cercle, la proposition 22.11 fonctionne partout en posant  $i_C(O) = \infty$  et  $i_C(\infty) = O$ .

De plus en pensant à la projection stéréographique, ce serait logique : quelle que soit la direction dans laquelle un point s'éloigne de  $z = 0$ , son image par l'inverse de la projection stéréographique s'approche du pôle nord.

### 23.6.1.2 Équation complexe d'un cercle

Un cercle de centre  $\omega \in \mathbb{C}$  et de rayon  $r$  a pour équation  $|z - \omega| = r$ , et nous avons les équivalences suivantes :

$$|z - \omega| = r \Leftrightarrow |z - \omega|^2 = r^2 \Leftrightarrow (z - \omega)(\bar{z} - \bar{\omega}) = r^2 \Leftrightarrow r\bar{r} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2. \quad (23.137)$$

Donc un cercle de centre  $\omega \in \mathbb{C}$  et de rayon  $r \in \mathbb{R}$  a pour équation

$$z\bar{z} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2. \quad (23.138)$$

#### Définition 23.63.

Un **cercle** dans  $\hat{\mathbb{C}}$  est une partie de la forme

$$C(\omega, r) = \{z \in \mathbb{C} \text{ tel que } z\bar{z} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2\} \quad (23.139)$$

avec  $\omega \in \mathbb{C}$  et  $r \in \mathbb{R}$ .

Dans la sphère de Riemann, aucun cercle ne contient le point  $\infty$ .

#### Exemple 23.64.

Trouvons le centre et le rayon du cercle d'équation

$$\bar{\omega}z + \omega\bar{z} = kz\bar{z} \quad (23.140)$$

avec  $k \neq 0$ . En divisant par  $k$  et en posant  $\sigma = \omega/k$  nous avons :

$$z\bar{z} - \bar{\sigma}z - \sigma\bar{z} = 0. \quad (23.141)$$

Cela est un cercle de centre  $\sigma$  et de rayon  $|\sigma|$ . En effet si  $z \in \mathbb{C}$  vérifie cette équation,

$$|z - \sigma|^2 = (z - \sigma)(\bar{z} - \bar{\sigma}) = \underbrace{z\bar{z} - \bar{\sigma}z - \sigma\bar{z}}_{=0} + |\sigma|^2 = |\sigma|^2, \quad (23.142)$$

c'est-à-dire que tous les points de  $\mathbb{C}$  qui vérifient l'équation donnée sont à la distance  $|\sigma|$  de  $\sigma$ . En particulier  $z = 0$  est sur le cercle.  $\triangle$

### 23.6.1.3 Cercle-droite

Une chose de bien avec les équations complexes, c'est que nous pouvons écrire les droites et les cercles avec le même type d'équations.

#### Lemme-Définition 23.65 ([678]).

Un **cercle-droite** est l'ensemble des points  $z \in \mathbb{C}$  tels que

$$az\bar{z} - \bar{\omega}z - \omega\bar{z} = k \quad (23.143)$$

avec  $a, k \in \mathbb{R}$  et  $\omega \in \mathbb{C}$ .

- (1) Si  $a = 0$ , cela est une droite ;
- (2) si  $a \neq 0$ , cela est un cercle.
- (3) Un cercle-droite peut être l'ensemble vide.

*Démonstration.* Si  $a = 0$  alors nous tombons tout de suite sur l'équation (23.135). Si  $a \neq 0$  alors nous pouvons diviser par  $a$ , poser  $\sigma = \omega/a$  et  $l = k/a$  pour obtenir

$$z\bar{z} - \bar{\sigma}z - \sigma\bar{z} = l, \quad (23.144)$$

qui est l'équation (23.138) d'un cercle ...ou pas tout à fait. En effet, (23.144) serait l'équation du cercle de centre  $\sigma$  et de rayon  $r$  donné par  $l = r^2 - |\sigma|^2$ , c'est à dire

$$r^2 = l + |\sigma|^2, \quad (23.145)$$

alors que rien n'assure que le nombre  $l + |\sigma|^2$  soit positif. Dans le cas où c'est positif, nous avons bien un cercle. Sinon c'est l'ensemble vide.  $\square$

REMooBMAEooHDvNID

**Remarque 23.66.**

Lorsque nous parlons de cercle-droite, nous parlons de partie de  $\hat{\mathbb{C}}$  et non de  $\hat{\mathbb{C}}$  parce que l'équation (23.143) a du mal à traiter le cas  $z = \infty$ . À cause du fait que nous avons décidé de donner le point  $\infty$  à toutes les droites, la fusion des notions de droites et de cercles n'est pas totale ; en tout cas pas en une seule équation.

EXooKFBlooOJKjGL

**Exemple 23.67** ([678]).

Soit le cercle de centre  $\omega = ir$  et de rayon  $r$ . Quelle que soit la valeur de  $r > 0$ , ce cercle passe par le point 0 et l'axe réel lui est tangent. L'équation de ce cercle est :

$$z\bar{r} + irz - ir\bar{z} = 0. \quad (23.146)$$

Vu que  $ir \neq 0$  nous pouvons diviser et obtenir

$$\frac{z\bar{z}}{ir} + z - \bar{z} = 0. \quad (23.147)$$

En faisant tendre  $r$  vers  $\infty$  nous obtenons  $z - \bar{z} = 0$ , c'est-à-dire l'équation de la droite réelle.

Cela explique pourquoi il est souvent dit qu'une droite est un cercle dont le rayon est à l'infini.

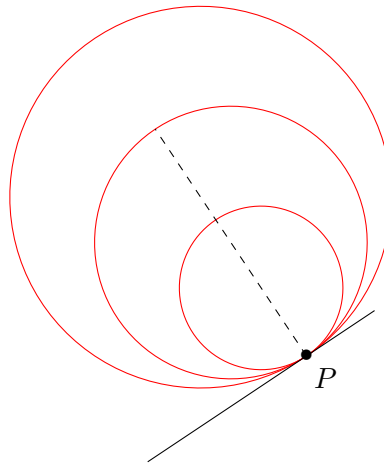
△

NORMooCXVJooMTMqEU

**23.68.**

Notons que l'exemple 23.67 est générique : prenez une droite  $\ell$ , un point  $P$  sur  $\ell$ , et considérez un cercle dont le centre est situé sur la perpendiculaire à  $\ell$  passant par  $P$ , et dont le rayon est tel que le cercle passe par  $P$ . En prenant  $|\omega - P| \rightarrow \infty$ , l'équation du cercle devient celle de la droite  $\ell$ .

Cela est particulièrement pratique lorsque nous travaillons dans  $\hat{\mathbb{C}}$  parce nous y avons une notion précise du point à l'infini. Notons que (peut-être contre-intuitivement), il existe un seul point à l'infini dans  $\hat{\mathbb{C}}$ . Et ce point est le centre de tous les cercles que l'on veut transformer en droites. Cela pose évidemment la question de savoir comment on définit précisément un cercle dont le centre est réellement  $\infty$ .

**23.6.1.4 Rotation-homothétie****Définition 23.69.**

Une **rotation-homothétie** est une application  $\hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$  de la forme  $z \mapsto \lambda z$  avec  $\lambda \in \mathbb{C}$ .

Le nom provient du fait que si  $\lambda$  est réel, alors  $z \mapsto \lambda z$  est une vraie homothétie, et si  $\lambda = e^{i\theta}$  alors  $z \mapsto e^{i\theta} z$  est une vraie rotation. Pour une valeur  $\lambda \in \mathbb{C}$  générique, l'application  $z \mapsto \lambda z$  est une composée des deux.

### 23.6.1.5 Application linéaire

SSUBSooRBCWooSCIQEL

Nous nous en voudrions de ne pas parler d'applications linéaires lorsque nous parlons de géométrie sur  $\hat{\mathbb{C}}$ . Soit  $\alpha \in \mathbb{C}^*$  et  $\beta \in \mathbb{C}$ . Lorsque nous parlons de l'application linéaire

$$\begin{aligned} f: \hat{\mathbb{C}} &\rightarrow \hat{\mathbb{C}} \\ z &\mapsto \alpha z + \beta, \end{aligned} \quad (23.148)$$

nous entendons implicitement que  $f(\infty) = \infty$ .

### 23.6.1.6 Inversion

SSUBSooPOUNooTPilbE

L'inversion d'un cercle de  $\mathbb{R}^2$  est définie par la proposition 22.5. De nombreuses propriétés y sont décrites, y compris son écriture complexe dans la proposition 22.10. Tout cela était du temps de  $\mathbb{R}^2$  ou de  $\mathbb{C}$ , mais maintenant nous sommes dans  $\hat{\mathbb{C}}$  et nous voulons plus.

DEFooIUTZooWRaXts

**Définition 23.70** ([678]).

Soient  $\omega \in \mathbb{C}$  et  $R \in \mathbb{R}^*$ . L'*inversion* de centre  $\omega$  et de *puissance*  $R^2$  est l'application

$$\begin{aligned} i: \hat{\mathbb{C}} &\rightarrow \hat{\mathbb{C}} \\ z &\mapsto \begin{cases} \frac{R^2}{\bar{z} - \bar{\omega}} + \omega & \text{si } z \in \mathbb{C} \setminus \{\omega\} \\ \infty & \text{si } z = \omega \\ \omega & \text{si } z = \infty. \end{cases} \end{aligned} \quad (23.149)$$

Notons que grâce aux conventions type  $1/0 = \infty$  et  $1/\infty = 0$ , nous pouvons nous contenter de la première formule pour tout  $z \in \hat{\mathbb{C}}$ , et nous n'avons en réalité pas besoin de décrire  $i(\infty)$  et  $i(\omega)$  séparément.

**Exemple 23.71.**

L'inversion de cercle de centre 0 et de rayon 1 est l'application  $z \mapsto \frac{1}{\bar{z}}$ , que l'on prolonge avec  $0 \mapsto \infty$  et  $\infty \mapsto 0$ .  $\triangle$

## 23.6.2 Homographies

La notion d'homographie est la définition 23.20. Pour une homographie  $\phi: P_1(\mathbb{C}) \rightarrow P_1(\mathbb{C})$  nous avons un isomorphisme d'espace vectoriel  $\bar{\phi}: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ . Vue la bijection (23.130), nous voulons plutôt travailler avec  $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$  qui est un ensemble avec lequel nous sommes plus familier. Nous allons donc travailler avec

$$\begin{aligned} \tilde{\phi}: \hat{\mathbb{C}} &\rightarrow \hat{\mathbb{C}} \\ \tilde{\phi} &= \varphi_0 \circ \phi \circ \varphi_0^{-1}. \end{aligned} \quad (23.150)$$

PROPooTZJBooPpowOo

**Proposition 23.72.**

L'application  $\phi = \varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0$  est une homographie de  $P(\mathbb{C}^2)$  si et seulement si l'application  $\tilde{\phi}: \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$  est de la forme

$$\tilde{\phi}(z) = \frac{az + b}{cz + d} \quad (23.151)$$

avec  $a, b, c, d \in \mathbb{C}$  tels que  $ad - bc \neq 0$ .

Par convention nous posons  $z/0 = \infty$  dès que  $z \neq 0$ ; en particulier  $\tilde{\phi}(\infty) = a/c$  et  $\tilde{\phi}(-d/c) = \infty$ .

*Démonstration.* Nous séparons la condition suffisante de la condition nécessaire.

(1)  $\Rightarrow$  La condition  $\pi \circ \tilde{\phi} = \phi \circ \pi$  (de (23.32)) nous dit que

$$[\tilde{\phi}(z_1, z_2)] = \phi([z_1, z_2]), \quad (23.152)$$

et comme  $\bar{\phi}$  est un isomorphisme d'espace vectoriel nous avons  $a, b, c, d \in \mathbb{C}$  vérifiant  $ad - cb \neq 0$  pour lesquels

$$\bar{\phi}(z_1, z_2) = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix}. \quad (23.153)$$

Soit  $z \in \mathbb{C}$ . Alors nous avons

$$\tilde{\phi}(z) = (\varphi_0 \circ \phi)[z, 1] = \varphi_0([\bar{\phi}(z, 1)]) = \varphi_0([az + b, cz + d]) = \frac{az + b}{cz + d}. \quad (23.154)$$

Il est important de comprendre que cette formule fonctionne pour tout  $z \in \mathbb{C}$ . En effet nous pourrions avoir un doute sur  $z = -d/c$ . D'abord si  $c = 0$  alors  $d \neq 0$  et ce problème n'existe pas : le dénominateur est toujours non nul. Nous avons donc seulement un doute lorsque  $c \neq 0$ . Dans ce cas,

$$\tilde{\phi}(-d/c) = \frac{-\frac{ad}{c} + b}{0}. \quad (23.155)$$

Mais  $c \neq 0$ , donc le numérateur est non nul. Or lorsque  $z \neq 0$  nous avons posé  $z/0 = \infty$ , donc dans notre cas,

$$\tilde{\phi}(-d/c) = \infty \quad (23.156)$$

automatiquement, et cela est encodé dans la formule (23.154).

Il nous reste à déterminer  $\tilde{\phi}(\infty)$ . Nous avons :

$$\tilde{\phi}(\infty) = (\varphi_0 \circ \phi)([1, 0]) = \varphi_0[\bar{\phi}(1, 1)] = \varphi_0([a, c]) = \begin{cases} a/c & \text{si } c \neq 0 \\ \infty & \text{si } c = 0 \end{cases} \quad (23.157)$$

où la distinction entre les deux cas n'est pas fondamentale parce que si  $c = 0$  alors  $a \neq 0$  et  $a/c = \infty$ .

(2)  $\Leftarrow$

En notant  $Z = [z_1, z_2]$  (pour  $z_1, z_2 \in \mathbb{C}$ ) nous avons

$$\phi([z_1, z_2]) = \varphi^{-1} \left( \frac{a\varphi_0(Z) + b}{c\varphi_0(Z) + d} \right) = [a\varphi_0(Z) + b, c\varphi_0(Z) + d]. \quad (23.158)$$

Nous définissons  $\bar{\phi}$  par son action sur les vecteurs de base :  $\bar{\phi}(1, 0) = (a, c)$  et  $\bar{\phi}(0, 1) = (b, d)$ . Nous avons bien l'isomorphisme d'espace vectoriel  $\bar{\phi}: \mathbb{C}^2 \rightarrow \mathbb{C}^2$  dont nous avons besoin pour la définition 23.20. D'abord le fait que  $ad - cb$  soit non nul assure que le  $\bar{\phi}$  ainsi défini est bien bijectif<sup>9</sup>. Et de plus ce  $\bar{\phi}$  vérifie la condition

$$\pi(\bar{\phi}(z_1, z_2)) = \pi((az_1, cz_1) + (bz_2, dz_2)) \quad (23.159a)$$

$$= [az_1 + bz_2, cz_1 + dz_2] \quad (23.159b)$$

$$= [a\varphi_0(Z) + b, c\varphi_0(Z) + d] \quad (23.159c)$$

$$= \phi(\pi(z_1, z_2)). \quad (23.159d)$$

Nous avons utilisé la notion de classe pour diviser par  $z_2$  et faire apparaître  $\varphi_0(Z)$ .

□

### Remarque 23.73.

En prenant les conventions relativement claires  $\infty \times a = \infty$  (pour  $a \neq 0$ ) et  $\infty \pm a = \infty$  (avec  $a \neq \infty$ ), alors tout est dans la formule

$$\bar{\phi}(z) = \frac{az + b}{cz + d} \quad (23.160)$$

avec  $ad - cb \neq 0$ . Il n'y a pas besoin de traiter séparément le cas  $z = \infty$  ou  $z = -d/c$ .

9. Si vous n'en êtes pas convaincu, écrivez la matrice de l'application qui envoie  $(1, 0)$  sur  $(a, c)$  et  $(0, 1)$  sur  $(b, d)$ , et demandez-vous sous quelle condition elle est inversible.



DEFooAMQHooPFUgIa

**Définition 23.74.**

Nous aimons tellement l'identification  $\varphi_0: P(\mathbb{C}^2) \rightarrow \hat{\mathbb{C}}$  que nous allons parler d'**homographie** sur  $\hat{\mathbb{C}}$  pour les applications  $\tilde{\phi}$  de la forme

$$\tilde{\phi}(z) = \frac{az + b}{cz + d} \quad (23.161)$$

avec  $ad - cb \neq 0$ .

NORMooCVYKooYvJleE

**23.75.**

La proposition 23.72 nous indique que les homographies de  $\hat{\mathbb{C}}$  sont de la forme  $\phi = \varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0$  pour une homographie  $\phi: P(\mathbb{C}^2) \rightarrow P(\mathbb{C}^2)$ .

PROFooSQFOooRginjJ

**Proposition 23.76** ([678]).

L'application  $h: \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$  associée à une homographie est soit linéaire, soit de la forme  $h = l_1 \circ \iota \circ l_2$  où  $l_i$  sont linéaires et  $\iota$  est l'application  $z \mapsto 1/z$ .

*Démonstration.* Commençons par une remarque : lorsque nous parlons d'une application linéaire, c'est au sens de la note 23.6.1.5 qui explique qu'une application linéaire sur  $\mathbb{C}$  est automatiquement prolongée à  $\hat{\mathbb{C}}$  par  $f(\infty) = \infty$ .

Soit donc l'application

$$h(z) = \frac{az + b}{cz + d}. \quad (23.162)$$

Si  $c = 0$ , alors c'est une application linéaire et la preuve est terminée. Nous supposons que  $c \neq 0$ . Nous posons

$$l_2(z) = cz + d, \quad (23.163)$$

et ensuite  $l_1(z) = \alpha z + \beta$  avec  $\alpha$  et  $\beta$  à déterminer. Un peu de calcul :

$$(l_1 \circ \iota \circ l_2)(z) = l_1\left(\frac{1}{cz + d}\right) = \frac{\alpha + \beta cz + \beta d}{cz + d}, \quad (23.164)$$

et en imposant que cela soit égal à  $\frac{az+b}{cz+d}$  nous trouvons  $\beta = a/c$  et  $\alpha = b - ad/c$ . Il est vite vérifié que ces choix donnent le bon résultat.  $\square$

NORMooMMKOooQlzzqJ

**23.77.**

Vu que les applications linéaires sont des composées d'une translation et d'une rotation-homothétie, et que l'application  $\iota$  est une composée d'une inversion  $z \mapsto 1/\bar{z}$  et d'une réflexion  $z \mapsto \bar{z}$ , toutes les homographies sont des composées des éléments suivants :

- inversion<sup>10</sup>  $z \mapsto 1/\bar{z}$ , prolongée par  $i(\infty) = 0$  et  $i(0) = \infty$  ;
- réflexion  $z \mapsto \bar{z}$  ;
- translation  $z \mapsto z + \alpha$  avec  $\alpha \in \mathbb{C}$  ;
- rotation-homothétie  $z \mapsto \lambda z$  avec  $\lambda \in \mathbb{C}$ .

Toutes ces opérations sont prolongées à  $\hat{\mathbb{C}}$  par  $1/\infty = 0$ ,  $\lambda \cdot \infty = \omega$  (si  $\lambda \neq 0$ ) et  $\infty + \lambda = \infty$ . Nous ne définissons pas  $0 \cdot \infty$  et  $\infty - \infty$ .

Certes nous pouvons construire des homographies à partir d'ingrédients dont la conjugaison complexe. Il ne faudrait cependant pas déduire que cette conjugaison est une homographie.

LEMooGDDJooBpJIUf

**Lemme 23.78.**

La conjugaison complexe n'est pas une homographie.

10. Oui, c'est l'inversion de la géométrie hyperbolique, voir 23.6.1.6.

*Démonstration.* Si elle l'était nous aurions des nombres  $a, b, c, d \in \mathbb{C}$  tels que  $ad - bc \neq 0$  et

$$\frac{az + b}{cz + d} = \bar{z} \quad (23.165)$$

pour tout  $z \in \mathbb{C}$ .

En posant  $z = 0$  nous avons déjà  $b/d = 0$ , c'est-à-dire  $b = 0$ . Avec  $z = 1$  nous trouvons alors  $a/(c + d) = 1$ , c'est-à-dire

$$a = c + d. \quad (23.166)$$

(1) Si  $ci + d \neq 0$

Dans ce cas nous pouvons évaluer (23.165) en  $z = i$  et avoir  $a = -ci + d$ . Mais comme nous avons déjà  $a = c + d$  nous déduisons  $c = 0$ . Nous restons donc avec

$$\frac{a}{d}z = \bar{z} \quad (23.167)$$

pour tout  $z$ . En prenant  $z = 1$  puis  $z = i$ , il est vite remarqué que cela n'est pas possible.

(2) Si  $ci + d = 0$

Nous rappelons que  $ad \neq 0$ . Nous écrivons l'équation avec  $z = -i$  pour trouver

$$\frac{-ai}{-ci + d} = i, \quad (23.168)$$

qui donne immédiatement  $a = ci - d$ . Nous avons donc les trois équations

$$\begin{cases} c = id & (23.169a) \\ a = ci - d & (23.169b) \\ a = c + d. & (23.169c) \end{cases}$$

Une tentative de résolution tombe rapidement sur une impossibilité (en substituant la première dans les deux autres et en comparant les deux valeurs de  $a$  par exemple).

□

La proposition suivante ressemble à s'y méprendre à la proposition 22.11, mais elle diffère en deux points. D'abord elle ne traite que de l'inversion par rapport à l'origine, mais surtout, elle traite le point  $z = \infty$ . C'est un avantage de travailler sur  $\hat{\mathbb{C}}$  plutôt que sur  $\mathbb{R}^2$ .

**Proposition 23.79** (Inversion de cercles et de droites).

*L'inversion dans  $\hat{\mathbb{C}}$  envoie*

- (1) *une droite passant par 0 sur elle-même*
- (2) *une droite ne passant pas par 0 sur un cercle passant par 0.*
- (3) *un cercle ne passant pas par 0 en un cercle ne passant pas par 0.*
- (4) *un cercle passant par 0 en une droite ne passant pas par 0.*

*Démonstration.* Décomposition en tous les cas possibles.

(1) Droite passant par 0

La façon la plus simple de traiter la droite passant par 0 est de l'écrire sous forme paramétrique :

$$z(t) = te^{i\theta} \quad (23.170)$$

pour  $\theta$  fixé et  $t \in \mathbb{R} \cup \{\infty\}$ . En appliquant l'inversion :

$$\iota(z(t)) = \frac{1}{(te^{i\theta})^*} = \frac{1}{t}e^{i\theta}. \quad (23.171)$$

Notons que les cas particuliers fonctionnent : pour  $t = 0$  nous avons le point  $\infty$  et pour  $t = \infty$  nous avons 0.

(2) **Droite ne passant pas par 0**

Une droite ne passant pas par  $z = 0$  est un ensemble de la forme

$$d(\omega, k) = \{z \in \mathbb{C}^* \text{ tel que } \bar{\omega}z + \omega\bar{z} = k\} \cup \{\infty\} \quad (23.172)$$

avec  $k \neq 0$ . Étant donné que  $\iota$  est une bijection et même une involution nous avons  $z \in \iota(d(\omega, k))$  si et seulement si  $\tau(z) \in d(\omega, k)$ . L'équation est donc, pour  $z \neq 0$  :

$$\frac{\bar{\omega}}{z} + \frac{\omega}{\bar{z}} = k. \quad (23.173)$$

Et comme  $z \neq 0$  nous pouvons multiplier par  $z\bar{z}$  pour trouver  $\bar{z}\bar{\omega} + z\omega = k z\bar{z}$ . Donc

$$\iota(d(\omega, k)) = \{z \in \mathbb{C}^* \text{ tel que } \bar{z}\bar{\omega} + z\omega = k z\bar{z}\} \cup \{0\}. \quad (23.174)$$

Dans l'ensemble, nous pouvons renommer  $z$  et  $\bar{z}$  pour avoir une forme plus symétrique. De plus il se fait que  $z = 0$  vérifie l'équation donnée; nous pouvons donc lever la condition  $z \in \mathbb{C}^*$  et ne plus ajouter  $\{0\}$  à côté :

$$\iota(d(\omega, k)) = \{z \in \mathbb{C} \text{ tel que } \bar{\omega}z + \omega\bar{z} = k z\bar{z}\}. \quad (23.175)$$

Cela est l'équation d'un cercle passant par l'origine (définition 23.63).

(3) **Cercle ne passant pas par 0**

Nous considérons le cercle  $C(\omega, r)$  avec  $r^2 \neq |\omega|^2$ . Il ne contient ni  $\infty$  ni 0 et nous avons alors

$$\iota(C(\omega, r)) = \{z \in \mathbb{C}^* \text{ tel que } \frac{1}{z\bar{z}} - \bar{\omega}\frac{1}{\bar{z}} - \omega\frac{1}{z} = r^2 - |\omega|^2\}. \quad (23.176)$$

Vu que  $z$  n'est jamais nul nous pouvons multiplier l'équation par  $z\bar{z}$  :

$$\iota(C(\omega, r)) = \{z \in \mathbb{C} \text{ tel que } (|\omega|^2 - r^2)z\bar{z} - \bar{\omega}z - \omega\bar{z} = -1\}. \quad (23.177)$$

Le coefficient  $|\omega|^2 - r^2$  est non nul par hypothèse et cet ensemble est un cercle par le lemme 23.65. Il ne passe manifestement pas par  $z = 0$ .

(4) **Cercle passant par 0**

Le cercle passe par 0, et donc son image par  $\infty$ . Nous écrivons alors

$$\iota(C(\omega, |\omega|)) = \iota(C(\omega, |\omega|) \setminus \{0\}) \cup \{\infty\}. \quad (23.178)$$

Nous avons

$$C(\omega, |\omega|) \setminus \{0\} = \{z \in \mathbb{C}^* \text{ tel que } z\bar{z} - \bar{\omega}z - \omega\bar{z} = 0\}, \quad (23.179)$$

et un calcul usuel donne

$$\iota(C(\omega, |\omega|) \setminus \{0\}) = \{z \in \mathbb{C}^* \text{ tel que } 1 - \bar{\omega}z - \omega\bar{z} = 0\}, \quad (23.180)$$

et donc

$$\iota(C(\omega, |\omega|)) = d(\omega, 1), \quad (23.181)$$

en nous souvenant que le point  $\infty$  est contenu dans  $d(\omega, 1)$ .

□

PROPooYFJBooAWxFIs

**Proposition 23.80.**

*Une homographie conserve l'ensemble des droites et cercles de  $\hat{\mathbb{C}}$ .*

Attention : cela ne veut pas dire qu'une homographie transforme une droite en une droite et un cercle en un cercle. Ça veut dire qu'une homographie transforme une droite en une droite ou un cercle et un cercle en une droite ou un cercle.

*Démonstration.* Nous savons par la proposition 22.11 et 23.77 que les homographies se décomposent en inversion, réflexion, translation et rotation-homothétie.

À part pour l'inversion, tout est clair comment ça fonctionne hein. En ce qui concerne l'inversion, nous avons la proposition 23.79 qui donne déjà toutes les réponses.  $\square$

### 23.81.

Les homographies préservent les angles, c'est l'objet du théorème suivant. Il ne faudrait cependant pas croire que si  $A$ ,  $B$  et  $C$  sont trois points distincts, l'angle entre  $\overline{AC}$  et  $\overline{BC}$  est le même que celui entre  $\overline{f(A)f(C)}$  et  $\overline{f(B)f(C)}$  dès que  $f$  est une homographie. Cela serait préserver les angles globalement, c'est-à-dire préserver les angles lorsque les points sont déplacés par  $f$ .

Nous allons regarder les angles locaux, c'est-à-dire lorsque les points sont déplacés par  $df$ .

### Définition 23.82.

Nous disons qu'une application  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  préserve localement les angles non orientés lorsque

$$\cos(df_a(u), df_a(v)) = \cos(u, v) \quad (23.182)$$

pour tout  $a \in \mathbb{R}^2$  et  $u, v \in \mathbb{R}^2$ . Ici il est mieux de penser à  $u, v \in T_a \mathbb{R}^2$  pour qui sait les espaces tangents en géométrie différentielle.

Voir la définition de l'angle 18.53.

### Théorème 23.83.

Les homographies de  $P(\mathbb{C}^2)$  préservent localement les angles non orientés.

*Démonstration.* En ce qui concerne les translations, dilatations et rotations, les choses sont claires. Vérifions pour l'inversion, qu'il faut interpréter dans  $\mathbb{R}^2$ . Pour  $z = x + iy$  nous avons

$$\iota(z) = \frac{1}{z} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2}. \quad (23.183)$$

Nous devons donc étudier la fonction

$$\begin{aligned} F: \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (x, y) &\mapsto \left( \frac{x}{r^2}, -\frac{y}{r^2} \right) \end{aligned} \quad (23.184)$$

où nous avons posé  $r^2 = x^2 + y^2$  pour simplifier les notations.

Soient deux vecteurs  $u, v \in \mathbb{R}^2$  et un point  $a \in \mathbb{R}^2$ . Nous devons prouver que

$$\frac{u \cdot v}{\|u\| \|v\|} = \frac{dF_a(u) \cdot dF_a(v)}{\|dF_a(u)\| \|dF_a(v)\|}. \quad (23.185)$$

Pour cela, nous pourrions calculer  $dF_a$  et passer en coordonnées polaires[675] mais nous préférons faire les calculs à la dure parce que nous avons Sage avec nous.

Nous notons  $A$  la matrice de  $dF$  en  $a = (x, y)$  et nous avons

$$Au \cdot Av = A^t Au \cdot v \quad (23.186)$$

ainsi que  $\|u\| = \sqrt{u \cdot u} = \sqrt{A^t Au \cdot u}$ , de telle sorte qu'il devienne urgent de calculer  $A^t A$ . Voici le calcul :

```

1 var('y')
2
3 # les fonctions coordonnées
4 F1(x,y)=x/(x**2+y**2)
5 F2(x,y)=-y/(x**2+y**2)
```

```

6
7 # La matrice différentielle :
8 A=matrix( [ [F1.diff(x).simplify_full(),F1.diff(y).simplify_full()←
    ()], [F2.diff(x).simplify_full(),F2.diff(y).simplify_full()] ←
    ] )
9
10 # Quelqu'un peut expliquer pourquoi ceci ne fonctionne pas ?
11 # A=matrix( [ [F1.diff(x),F1.diff(y)], [F2.diff(x),F2.diff(y)] ←
    ].simplify_full()
12
13 # A^tA
14 S=A.transpose()*A
15 S=S.simplify_full() # Mais ça, ça marche !!
16 print(S)

```

tex/sage/sageSnip009.sage

Le résultat est que

$$A^t A = \begin{pmatrix} \frac{1}{r^4} & 0 \\ 0 & \frac{1}{r^4} \end{pmatrix} = \frac{1}{r^4} \text{Id}. \quad (23.187)$$

La vérification de (23.185) est alors immédiate. □

### 23.6.3 Birapport

Nous introduisons maintenant quelque chose qui s'appelle le « birapport » et qui n'est à priori pas du tout lié au birapport défini en 23.42.

DEFooQYHVooMZwQMB

**Définition 23.84** (Birapport dans  $\hat{\mathbb{C}}$ [605]).

Soient  $a, b, c, x \in \hat{\mathbb{C}}$  où  $a, b$  et  $c$  sont distincts. Le **birapport** de ces quatre nombres est l'élément de  $\hat{\mathbb{C}}$  donné par, si  $a, b, c \neq \infty$  :

$$[a, b, c, x] = \frac{(a-c)(b-x)}{(b-c)(a-x)}, \quad (23.188) \quad \text{EQooQJWZooOXKslh}$$

et

$$[\infty, b, c, x] = \frac{b-x}{b-c} \quad (23.189a) \quad \text{SUBEQooNSONooYUhuZB}$$

$$[a, \infty, c, x] = \frac{a-c}{a-x} \quad (23.189b)$$

$$[a, b, \infty, x] = \frac{b-x}{a-x} \quad (23.189c)$$

#### 23.85.

Notons la « logique » des cas particuliers. Pour le premier, si  $a \rightarrow \infty$  tandis que les autres restent dans  $\mathbb{C}$  alors  $a-c$  et  $a-x$  deviennent du même ordre de grandeur et se simplifient. Il reste les deux autres parties de la fraction.

C'est cette même logique qui, partant de  $[a, b, \infty, x] = \frac{b-x}{a-x}$  donne

$$[a, b, \infty, \infty] = 1 \quad (23.190)$$

comme il se doit si nous avons l'intention de ressembler au lemme 23.47.

L'objet « birapport » introduit ici est évidemment lié au birapport sur  $P(\mathbb{C}^2)$  défini plus haut. Le lien est la proposition suivante.

PROPooLKQZooEOrijwC

**Proposition 23.86.**

Soit l'application  $\varphi_0: P(\mathbb{C}^2) \rightarrow \hat{\mathbb{C}}$  définie en 23.36. Si  $A, B, C, X \in P(\mathbb{C}^2)$  alors

$$[A, B, C, X]_{\varphi_0} = [\varphi_0(A), \varphi_0(B), \varphi_0(C), \varphi_0(X)]. \quad (23.191)$$

EQooEOZZooMRHJfH

Cela est une égalité dans  $\hat{\mathbb{C}}$ .

*Démonstration.* Nous écrivons  $A = [a_1, a_2]$ ,  $B = [b_1, b_2]$ ,  $C = [c_1, c_2]$  avec  $a_1, a_2, b_1, b_2, c_1, c_2 \in \mathbb{C}$ . Par définition  $[A, B, C, X]_{\varphi_0} = (\varphi_0 \circ \phi)(X)$  où  $\phi: P(\mathbb{C}^2) \rightarrow P(\mathbb{C}^2)$  est l'unique homographie telle que

SUBEQSooTFWTooRnijrY

$$\phi[a_1, a_2] = [1, 0] \quad (23.192a)$$

$$\phi[b_1, b_2] = [0, 1] \quad (23.192b)$$

$$\phi[c_1, c_2] = [1, 1] \quad (23.192c)$$

Une des difficultés de cette preuve va être de calculer ce  $\phi$ . D'abord nous pouvons introduire  $\tilde{\phi} = \varphi_0 \circ \phi \circ \varphi_0^{-1}$  qui est obligatoirement (proposition 23.72) de la forme

$$\tilde{\phi}(z) = \frac{\alpha z + \beta}{\gamma z + \delta}. \quad (23.193)$$

Nous allons imposer les relations (23.192) pour déterminer les coefficients  $\alpha, \beta, \gamma$  et  $\delta$ .

D'abord

$$(\varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0)([a_1, a_2]) = (\varphi_0^{-1} \circ \tilde{\phi})(a_1/a_2) \quad (23.194a)$$

$$= \varphi_0^{-1} \left( \frac{\alpha \frac{a_1}{a_2} + \beta}{\gamma \frac{a_1}{a_2} + \delta} \right) \quad (23.194b)$$

$$= \left[ \alpha \frac{a_1}{a_2} + \beta, \gamma \frac{a_1}{a_2} + \delta \right]. \quad (23.194c)$$

Égaliser cela à  $[1, 0]$  donne

$$\begin{cases} \alpha \frac{a_1}{a_2} + \beta \neq 0 \\ \gamma \frac{a_1}{a_2} + \delta = 0. \end{cases} \quad (23.195a)$$

$$\quad (23.195b)$$

Donc nous avons déjà

$$\tilde{\phi}(z) = \frac{\alpha z + \beta}{\gamma(z - \frac{a_1}{a_2})} = \frac{\alpha z + \beta}{\gamma(z - \varphi_0(A))}. \quad (23.196)$$

En y imposant la contrainte  $(\varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0)([b_1, b_2]) = [0, 1]$  nous trouvons les contraintes

$$\begin{cases} \gamma(\varphi_0(B) - \varphi_0(A)) \neq 0 \\ \beta = -\alpha \varphi_0(B). \end{cases} \quad (23.197a)$$

$$\quad (23.197b)$$

Nous avons décidé d'écrire  $\varphi_0(A)$  au lieu de  $a_1/a_2$  à la fois pour un soucis de simplification d'écriture et dans le but de ressembler à (23.191). En substituant :

$$\tilde{\phi}(z) = \frac{\alpha(z - \varphi_0(B))}{\gamma(z - \varphi_0(A))}. \quad (23.198)$$

La condition pour  $[c_1, c_2]$  donne

$$[\alpha(\varphi_0(C) - \varphi_0(B)), \gamma(\varphi_0(C) - \varphi_0(A))] = [1, 1], \quad (23.199)$$

ce qui donne

$$\alpha(\varphi_0(C) - \varphi_0(B)) = \gamma(\varphi_0(C) - \varphi_0(A)). \quad (23.200)$$

Nous avons alors

$$\gamma = \alpha \frac{\varphi_0(C) - \varphi_0(B)}{\varphi_0(C) - \varphi_0(A)}, \quad (23.201)$$

et les  $\alpha$  se simplifient dans la formule pour  $\tilde{\phi}$  :

$$\tilde{\phi}(z) = \frac{(z - \varphi_0(B))(\varphi_0(C) - \varphi_0(A))}{(z - \varphi_0(A))(\varphi_0(C) - \varphi_0(B))}. \quad (23.202)$$

Par la proposition 23.72, l'application  $\varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0$  est une homographie. Nous pouvons donc calmement calculer le birapport  $[A, B, C, X]_{\varphi_0}$  de la façon suivante :

$$[A, B, C, X]_{\varphi_0} = (\varphi_0 \circ \phi)(X) \quad (23.203a)$$

$$= (\varphi_0 \circ \varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0)(X) \quad (23.203b)$$

$$= \tilde{\phi}(\varphi_0(X)) \quad (23.203c)$$

$$= \frac{(\varphi_0(X) - \varphi_0(B))(\varphi_0(C) - \varphi_0(A))}{(\varphi_0(X) - \varphi_0(A))(\varphi_0(C) - \varphi_0(B))} \quad (23.203d)$$

$$= [\varphi_0(A), \varphi_0(B), \varphi_0(C), \varphi_0(X)]. \quad (23.203e)$$

□

PROF00QGPF00ReNaGq

### Proposition 23.87.

Les homographies de  $\hat{\mathbb{C}}$  conservent le birapport.

*Démonstration.* Ici le mot « homographie » réfère à la définition 23.74 et le birapport à 23.84. Soient  $a, b, c, x \in \hat{\mathbb{C}}$  et une homographie  $\tilde{\phi}: \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$ . Il existe une homographie  $\phi: P(\mathbb{C}^2) \rightarrow P(\mathbb{C}^2)$  telle que  $\tilde{\phi} = \varphi_0 \circ \phi \varphi_0^{-1}$ . Alors

$$[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)] = [(\phi \circ \varphi_0^{-1})(a), (\phi \circ \varphi_0^{-1})(b), (\phi \circ \varphi_0^{-1})(c), (\phi \circ \varphi_0^{-1})(x)]_{\varphi_0} \quad (23.204a)$$

$$= [\varphi_0^{-1}(a), \varphi_0^{-1}(b), \varphi_0^{-1}(c), \varphi_0^{-1}(x)]_{\varphi_0} \quad (23.204b)$$

$$= [a, b, c, x]. \quad (23.204c)$$

Justifications :

- Identification des birapports sur  $P(\mathbb{C}^2)$  et sur  $\hat{\mathbb{C}}$ , proposition 23.86.
- Invariance du birapport sous les homographies (dans  $P(\mathbb{C}^2)$ ), proposition 23.49.

□

PROF00SGCJooLnOLCx

### Proposition 23.88.

Soient des points  $a, b, c, x$  dans  $\hat{\mathbb{C}}$  avec  $a, b, c$  distincts. Ils sont alignés ou cocycliques si et seulement si  $[a, b, c, x] \in \mathbb{R}$ .

*Démonstration.* Nous allons faire plusieurs cas. Mais dans tous les cas vous pouvez relire la définition des angles orientés 18.146 et la partie sur les angles dans les nombres complexes 18.10.11.

#### (1) Tous les points sont distincts et dans $\mathbb{C}$

D'une part, nous savons que le nombre complexe  $re^{i\theta}$  est réel si et seulement si  $\theta \in [0]_{\pi}$ , et d'autre part l'argument du birapport (23.188) est

$$[\vec{ca}, \vec{cb}] + [\vec{xb}, \vec{xa}]. \quad (23.205)$$

Le birapport est réel si et seulement si

$$[\vec{ca}, \vec{cb}] - [\vec{xb}, \vec{xa}] \in [0]_{\pi}. \quad (23.206)$$

À gauche nous avons une classe modulo  $2\pi$  et à droite une classe modulo  $\pi$ . L'égalité signifie qu'il y a un représentant du membre de gauche qui appartient au membre de droite. Si vous

aimez faire très attention à ce que signifient les notations, voici trois manières d'écrire la condition, par ordre croissant de précision :

$$[\vec{ca}, \vec{cb}] - [\vec{xa}, \vec{xb}] = [0]_\pi, \quad (23.207a)$$

$$[\vec{ca}, \vec{cb}] - [\vec{xa}, \vec{xb}] \in [0]_\pi, \quad (23.207b)$$

$$[\vec{ca}, \vec{cb}] - [\vec{xa}, \vec{xb}] \subset [0]_\pi. \quad (23.207c)$$

Nous avons donc que le birapport est réel si et seulement si la condition (23.207c) est vérifiée. D'après le théorème 18.154, cette dernière condition est équivalente à dire que les points  $a$ ,  $b$ ,  $c$  et  $x$  sont alignés.

## (2) Pas quatre points distincts, dans $\mathbb{C}$

Nous supposons encore que  $a$ ,  $b$ ,  $c$  et  $x$  sont dans  $\mathbb{C}$ . Mais nous supposons que  $x$  est un de  $a$ ,  $b$  ou  $c$ . Vu que par hypothèse  $a$ ,  $b$  et  $c$  sont distincts, c'est le seul cas à considérer dans la catégorie des 4 points non distincts.

Trois points sont toujours alignés ou cocycliques<sup>11</sup>. Donc nous devons seulement montrer que dans ce cas le birapport est toujours dans  $\hat{\mathbb{R}}$ . Par définition,

- Si  $x = a$  alors  $[a, b, c, x] = \infty$ ,
- Si  $x = b$  alors  $[a, b, c, x] = 0$ ,
- Si  $a = c$  alors  $[a, b, c, x] = 1$ .

Dans tous les cas de figure le birapport est dans  $\hat{\mathbb{R}}$ .

À ce niveau de la preuve nous devons encore vérifier les cas où  $a$ ,  $b$ ,  $c$  ou  $x$  valent  $\infty$ . Si l'un de  $a$ ,  $b$  ou  $c$  est  $\infty$  et si  $x$  l'est aussi, alors, comme  $\infty$  est aligné avec tout, nous avons seulement une droite passant par deux points. Il nous faut donc seulement regarder les cas où un seul des 4 points est  $\infty$ .

### (1) Si $a = \infty$ Le birapport est alors (par (23.189a))

$$[\infty, b, c, x] = \frac{b - x}{a - x}, \quad (23.208)$$

qui est un nombre à priori complexe dont le dénominateur est supposé non nul parce que le cas  $a = x$  est déjà traité. L'argument de ce nombre est dans la classe de l'angle orienté

$$\arg\left(\frac{b - x}{a - x}\right) \in [\vec{xa}, \vec{xb}]. \quad (23.209)$$

Le birapport est réel si et seulement si le membre de gauche est dans  $[0]_\pi$ . Et cela est justement le cas où le membre de droite donne des points alignés.

Les cas  $b = \infty$ ,  $c = \infty$  et  $x = \infty$  se traitent de la même manière. □

## 23.6.4 Division harmonique

### Définition 23.89.

Nous disons que les éléments  $a$ ,  $b$ ,  $c$  et  $x$  de  $\hat{\mathbb{C}}$  sont en **division harmonique** lorsque  $[a, b, c, x] = -1$ .

### 23.90.

Une chose qui sera utile par la suite est de remarquer que  $[a, b, c, \infty] = -1$  lorsque  $c = \frac{a+b}{2}$ .

Nous allons maintenant voir comment, pour  $a, b, c \in \mathbb{C}$  donnés nous pouvons construire  $x$  tels que  $a, b, c, x$  soient en division harmonique. Vu que trois points sont soit cocycliques soit alignés nous divisons la construction en deux parties.

Notons que si nous trouvons une construction qui donne un point  $x$  vérifiant  $[a, b, c, x] = -1$  alors nous prouvons au passage que la construction ne dépend pas des choix intermédiaires parce que il n'existe qu'un unique  $x$  tel que  $[a, b, c, x] = -1$  lorsque  $a, b, c$  sont donnés.

11. Si ils ne sont pas alignés, prendre la médiatrice du segment  $[a, b]$  et celle de  $[b, c]$ , et l'intersection vous donnera le centre d'un cercle passant par  $a$ ,  $b$  et  $c$ .



LEMooAEDTooKsUoPw

**Lemme 23.91** ([679, 680]).

Soient  $a, b, c$  cocycliques dans  $\mathbb{C}$ . Nous nommons  $\mathcal{C}$  le cercle contenant  $a, b$  et  $c$  ainsi que  $T_a$  et  $T_b$  les tangentes à  $\mathcal{C}$  en  $a$  et  $b$ . Soit  $m = T_a \cap T_b$  et la droite  $L = (mc)$ . Alors le point

$$x = (mc) \cap \mathcal{C} \quad (23.210)$$

vérifie  $[a, b, c, x] = -1$ .

Si  $m = \infty$  (arrive lorsque  $T_a \parallel T_b$ ) alors en guise de  $L$  nous prenons la parallèle à  $T_a$  passant par  $c$ .

*Démonstration.* Nous séparons les cas suivant que  $m = \infty$  ou non.

(1)  $m = \infty$

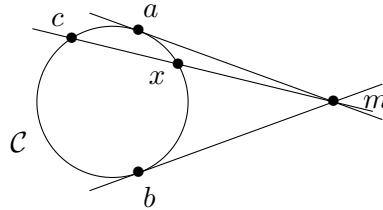
Les tangentes à  $\mathcal{C}$  en  $a$  et en  $b$  sont parallèles, c'est-à-dire que ces points sont diamétralement opposés sur  $\mathcal{C}$ . Les homographies préservent le birapport (proposition 23.87), et les rotations, dilatations et translations sont des homographies (voir 23.77).

Nous pouvons donc nous ramener au cas où  $\mathcal{C}$  est centré en 0 et de rayon 1 avec  $a = i$  et  $b = -i$ . Dans ce cas,  $c = e^{i\theta}$ . Vu que  $x$  est donné par l'intersection entre le cercle et la droite horizontale passant par  $c$  nous avons  $x = e^{i(\pi-\theta)}$ . Le birapport se calcule explicitement :

$$[a, b, c, x] = \frac{(i - e^{i\theta})(-i - e^{i(\pi-\theta)})}{(-i - e^{i\theta})(i - e^{i(\pi-\theta)})} = -1. \quad (23.211)$$

(2)  $m \neq \infty$

Nous sommes dans la situation suivante où à une translation près nous supposons  $x = 0$  :



Nous allons prouver que dans ce cas,  $[a, b, c, x] = -1$ . Pour cela nous considérons l'inversion de centre  $x$  (qui est  $x = 0$  par translation). Soit  $\tilde{\phi}$  cette homographie. Elle conserve le birapport, il nous allons voir que calculer  $[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)]$  se révèle être plus facile<sup>12</sup>.

Nous nommons  $A = (am)$ ,  $B = (bm)$ ,  $C = (cm)$  et  $\mathcal{C}$ , le cercle. Nous allons maintenant faire intensément usage de la proposition 23.79. Nous avons :

- $\tilde{\phi}(A)$  est un cercle passant par 0.
- $\tilde{\phi}(B)$  est un cercle passant par 0.
- $\tilde{\phi}(C)$  est la droite  $C$ .
- $\tilde{\phi}(\mathcal{C})$  est une droite ne passant pas par 0.

Les droites  $A$  et  $B$  se coupent en  $m$  et en  $\infty$  (qui sont des points distincts). Donc les cercles  $\tilde{\phi}(A)$  et  $\tilde{\phi}(B)$  se coupent en 0 et  $\tilde{\phi}(m)$ , aucun de ces deux points n'est sur la droite  $\tilde{\phi}(C)$ .

Par tangence, la droite  $A$  et le cercle  $\mathcal{C}$  se coupent en un seul point ( $a$ ). Donc  $\tilde{\phi}(A)$  coupe  $\tilde{\phi}(C)$  en un seul point,  $\tilde{\phi}(a)$ . Idem pour le cercle  $\tilde{\phi}(B)$ .

Nous avons donc que les cercles  $\tilde{\phi}(A)$  et  $\tilde{\phi}(B)$  sont tangents à la droite  $\tilde{\phi}(C)$  et se coupent en exactement deux points distincts (qui sont donc du même côté de la droite).

Nous nous intéressons à la droite  $\tilde{\phi}(C)$ . C'est une droite parce que c'est l'image d'une droite passant par 0. Elle passe par 0, par  $\tilde{\phi}(c)$  et  $\tilde{\phi}(m)$ . Le fait qu'elle passe par 0 et  $\tilde{\phi}(m)$  fait

12. Si vous n'avez peur d'aucun calculs, il suffit de poser  $a = e^{i\theta}$ ,  $b = e^{-i\theta}$  et  $c = e^{i\sigma}$  et vous êtes théoriquement capable de calculer les coordonnées de tous les points, y compris de  $x$  en termes de  $\theta$  et  $\sigma$ . Ensuite le calcul du birapport est explicite.

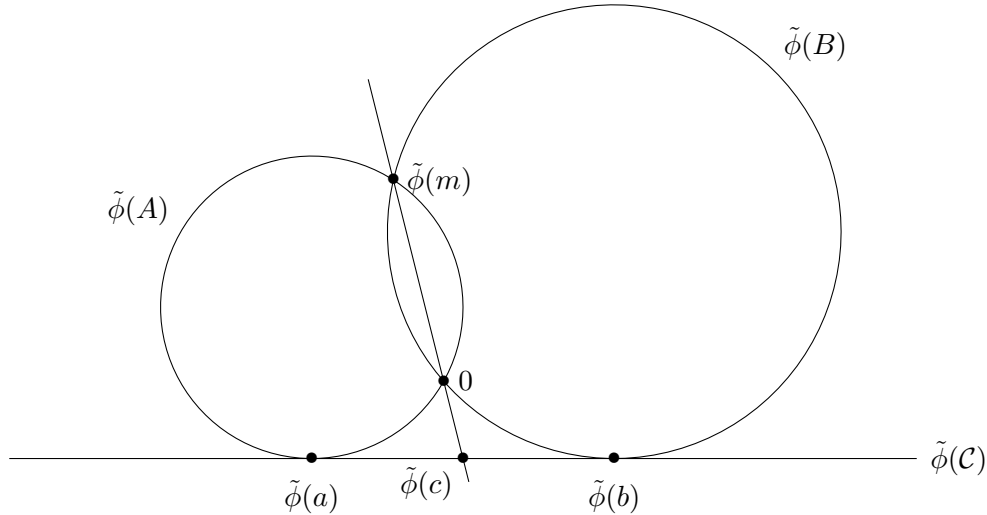
que c'est la droite passant par les deux intersections des cercles. Vu que  $c \in \mathcal{C} \cap C$ , le point d'intersection  $\tilde{\phi}(C) \cap \tilde{\phi}(C)$  est  $\tilde{\phi}(c)$ .

Quelle est la puissance du point  $\tilde{\phi}(c)$  par rapport au cercle  $\tilde{\phi}(A)$ ? En la calculant avec la droite  $\tilde{\phi}(C)$ , qui intersecte les deux cercles aux points déjà étudiés, la puissance est :

$$k = d(\tilde{\phi}(c), 0) \times d(\tilde{\phi}(c), \tilde{\phi}(m)). \quad (23.212)$$

Vu que ces points sont également sur le cercle  $\tilde{\phi}(B)$ , la puissance de  $\tilde{\phi}(c)$  par rapport à ce second point est la même.

Tout cela justifie le dessin suivant <sup>13</sup> :



Mais la droite passant par  $\tilde{\phi}(a)$  et  $\tilde{\phi}(c)$  (qui est tangente au cercle) permet également de calculer cette puissance :

$$k = d(\tilde{\phi}(a), \tilde{\phi}(c)). \quad (23.213)$$

Idem pour la puissance par rapport à l'autre cercle :

$$k = d(\tilde{\phi}(b), \tilde{\phi}(c)). \quad (23.214)$$

Nous en déduisons que  $\tilde{\phi}(c)$  est le milieu entre  $\tilde{\phi}(a)$  et  $\tilde{\phi}(b)$ .

Du coup

$$[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)] = [\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \infty] = -1 \quad (23.215)$$

en vertu de ce que nous avons raconté en [23.90](#).

□

LEMooYBTHooABWkeo

### Lemme 23.92 ([680]).

Soit  $a, b, c \in \mathbb{C}$  colinéaires. Soit  $m$  un point hors de cette droite. Nous considérons une droite issue de  $c$  coupant  $[ma]$  en  $p$  et  $[mb]$  en  $q$ .

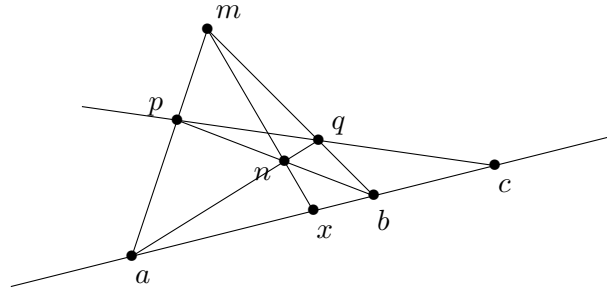
Nous construisons  $n = (aq) \cap (pb)$  et finalement  $x = (mn) \cap (ab)$ .

À la fin nous avons

$$[a, b, c, x] = -1. \quad (23.216)$$

*Démonstration.* Commençons par un dessin de la situation :

13. Bien que ce ne soit pas strictement nécessaire à la preuve, est-ce que vous savez si les deux cercles ont le même rayon? Et si par hasard la droite  $(\tilde{\phi}(m)\tilde{\phi}(c))$  n'arrive pas perpendiculairement à  $\tilde{\phi}(C)$ ?



Les points  $a$ ,  $b$  et  $m$  ne sont pas alignés, et nous pouvons les utiliser comme repère barycentrique (voir 11.51 pour savoir en deux mots ce que c'est). Nous nommons  $(\alpha, \beta, \gamma)$  les coordonnées de  $n$  dans ce système, c'est-à-dire que

$$\alpha \vec{na} + \beta \vec{nb} + \gamma \vec{nm} = 0. \quad (23.217)$$

Dans notre contexte, nous pouvons voir le vecteur  $\vec{st}$  comme une façon d'écrire le nombre  $t - s$ . Par la proposition 11.58 nous savons les coordonnées barycentriques de  $p$ ,  $x$  et  $q$  en regardant le triangle  $acb$ . Voici les coordonnées et les relations qu'elles signifient :

$$n = (\alpha, \beta, \gamma), \quad \alpha \vec{na} + \beta \vec{nb} + \gamma \vec{nm} = 0 \quad (23.218a)$$

$$p = (\alpha, 0, \gamma), \quad \alpha \vec{pa} + \gamma \vec{pm} = 0 \quad (23.218b)$$

$$q = (0, \beta, \gamma), \quad \beta \vec{qb} + \gamma \vec{qm} = 0 \quad (23.218c)$$

$$x = (\alpha, \beta, 0). \quad \alpha \vec{xa} + \beta \vec{xb} = 0. \quad (23.218d)$$

Nous voudrions maintenant voir les coordonnées de  $c$ . Nous posons  $c = (\lambda, \mu, \sigma)$  :

$$\lambda \vec{ca} + \mu \vec{cb} + \sigma \vec{cm} = 0. \quad (23.219)$$

Mais  $a$ ,  $b$  et  $c$  sont alignés, donc  $\vec{ca}$  et  $\vec{cb}$  sont colinéaires, alors que  $\vec{cm}$  n'est pas aligné avec les deux autres. L'annulation (23.219) demande donc l'annulation séparément

$$\begin{cases} \sigma \vec{cm} = 0 \\ \lambda \vec{ca} + \mu \vec{cb} = 0. \end{cases} \quad (23.220a)$$

$$\lambda \vec{ca} + \mu \vec{cb} = 0. \quad (23.220b)$$

Nous en déduisons que  $\sigma = 0$  et aussi que  $\lambda$  et  $\mu$  ne sont pas nuls. Nous posons arbitrairement  $\lambda = 1$  parce que les coordonnées barycentriques sont définies à coefficient multiplicatif près.

Nous imposons à présent le fait que  $p$ ,  $q$  et  $c$  sont alignés. Pour cela nous devons faire apparaître les vecteurs  $\vec{pq}$ ,  $\vec{pc}$ ,  $\vec{qc}$ . Vu le dessin et les relations disponibles (23.218) le mieux est d'utiliser les relations de Chasles (proposition 11.5) pour faire  $\vec{ca} = \vec{cp} + \vec{pa}$  et  $\vec{cb} = \vec{cq} + \vec{qb}$ . La relation (23.220b) devient :

$$\vec{cp} + \vec{pa} + \mu(\vec{cq} + \vec{qb}) = 0. \quad (23.221)$$

Les vecteurs  $\vec{cp}$  et  $\vec{cq}$  sont alignés, donc nous les écrivons ensemble. Les vecteurs  $\vec{pa}$  et  $\vec{qb}$  se transforment en utilisant les relations (23.218) :

$$\vec{cp} + \mu \vec{cq} - \frac{\gamma}{\alpha} \vec{pm} - \mu \frac{\gamma}{\beta} \vec{qm}. \quad (23.222)$$

Enfin nous voulons faire la somme du terme  $\vec{pm}$  avec le terme  $\vec{qm}$ . D'abord on change le signe :

$$\vec{cp} + \mu \vec{cq} - \frac{\gamma}{\alpha} \vec{pm} + \frac{\mu\gamma}{\beta} \vec{mq} \quad (23.223)$$

ensuite nous écrivons

$$\frac{\mu\gamma}{\beta} = \frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} - \frac{\gamma}{\alpha}, \quad (23.224)$$

et

$$\vec{cp} + \mu \vec{cq} - \frac{\gamma}{\alpha} (\vec{pm} + \vec{mq}) + \left( \frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} \right) \vec{mq} = 0. \quad (23.225)$$

Tout cela pour

$$\overrightarrow{cp} + \mu \overrightarrow{cq} - \frac{\gamma}{\alpha} \overrightarrow{pq} + \left( \frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} \right) \overrightarrow{mq} = 0. \quad (23.226)$$

Le dernier terme n'est pas colinéaire aux deux premiers et s'annule donc séparément :

$$\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} = 0. \quad (23.227)$$

Cela donne  $\mu = -\beta/\alpha$ .

Au final nous avons

$$\overrightarrow{ca} - \frac{\beta}{\alpha} \overrightarrow{cb} = 0 \quad (23.228)$$

et donc

$$\alpha \overrightarrow{ca} - \beta \overrightarrow{cb} = 0, \quad (23.229)$$

ce qui donne les coordonnées  $(\alpha, -\beta, 0)$  pour le point  $c$ .

Vu que nous sommes dans l'espace vectoriel  $\mathbb{C}$ , ce que nous notons  $\overrightarrow{AB}$  n'est rien d'autre que la différence  $B - A$  dans  $\mathbb{C}$ <sup>14</sup>. La relation (23.229) signifie donc

$$\alpha(a - c) = \beta(b - c). \quad (23.230)$$

Nous avons alors :

$$\frac{a - c}{b - c} = \frac{\beta}{\alpha}. \quad (23.231)$$

Par ailleurs, la relation (23.218d) à propos des coordonnées de  $x$  donne

$$\frac{a - x}{b - x} = -\frac{\beta}{\alpha}. \quad (23.232)$$

En égalisant les deux valeurs de  $\beta/\alpha$  nous trouvons :

$$\frac{a - c}{b - c} = \frac{x - a}{b - x}, \quad (23.233)$$

ce qui donne (via un petit jeu de signes)

$$\frac{(c - a)(x - b)}{(c - b)(x - a)} = -1. \quad (23.234)$$

C'est cela que nous voulions. □

### 23.6.5 Groupe circulaire

Nous avons vu que les homographies préservent l'ensemble des cercles et droites. Nous pouvons nous demander quel est le groupe maximum préservant l'ensemble des cercles et droites.

#### Définition 23.93.

Le **groupe circulaire** de  $\mathbb{C}$  est le groupe de transformations de  $\hat{\mathbb{C}}$  engendré<sup>15</sup> par les homographies<sup>16</sup> et la conjugaison complexe. Le groupe circulaire de l'espace projectif est l'ensemble des applications de la forme  $\varphi_0^{-1} \circ f \circ \varphi_0$  où  $f$  est un élément du groupe circulaire de  $\mathbb{C}$ .

Vu le lemme 23.78, la conjugaison complexe n'est pas une homographie. Donc cette définition n'est pas stupide : le groupe circulaire est strictement plus grand que le groupe des homographies.

14. Les mauvaises langues diront que tout le chapitre sur les espaces affines, et surtout la partie sur les barycentres ne sont rien d'autres que le snobisme d'écrire  $\overrightarrow{xy}$  au lieu de  $y - x$ . C'est aussi une facilité d'écriture.

15. Définition 1.420.

16. Homographie de  $\hat{\mathbb{C}}$  : définition 23.74.

LEMooOPOMooWZLSeH

**Lemme 23.94.**

Soit une application  $\alpha: \mathbb{C} \rightarrow \mathbb{C}$  fixant 1 et 0 et préservant les divisions harmoniques (c'est-à-dire tel que son prolongement à  $\hat{\mathbb{C}}$  donné par  $\alpha(\infty) = \infty$  préserve les divisions harmoniques). Alors  $\alpha$  est un automorphisme de corps<sup>17</sup>.

*Démonstration.* Nous savons que si  $a, b, c \in \mathbb{C}$  nous avons  $c = (a + b)/2$  si et seulement si  $[a, b, c, \infty] = -1$ . Vu que  $\alpha$  préserve les divisions harmoniques nous avons équivalence entre les affirmations suivantes :

$$c = \frac{a + b}{2} \quad (23.235a)$$

$$[a, b, c, \infty] = -1 \quad (23.235b)$$

$$[\alpha(a), \alpha(b), \alpha(c), \alpha(\infty)] = -1 \quad (23.235c)$$

$$\frac{\alpha(a) + \alpha(b)}{2} = \alpha(c). \quad (23.235d)$$

Donc  $\alpha$  préserve les milieux : pour tout  $a, b \in \mathbb{C}$  nous avons

$$\alpha\left(\frac{a + b}{2}\right) = \frac{\alpha(a) + \alpha(b)}{2}. \quad (23.236) \quad \text{EQooZCWFooZkwVYVW}$$

En particulier, cette relation avec  $b = 0$  donne (parce que  $\alpha(0) = 0$ ) :  $\alpha(a/2) = \alpha(a)/2$ . Nous avons au final, en utilisant cela en conjonction avec (23.236) :

$$\frac{\alpha(a) + \alpha(b)}{2} = \alpha\left(\frac{a + b}{2}\right) = \frac{\alpha(a + b)}{2}. \quad (23.237)$$

Cela démontre déjà que

$$\alpha(a + b) = \alpha(a) + \alpha(b). \quad (23.238)$$

En particulier  $\alpha(-a) = \alpha(0 - a) = \alpha(0) - \alpha(a) = -\alpha(a)$ .

Nous passons maintenant à la démonstration du fait que  $\alpha(ab) = \alpha(a)\alpha(b)$ . Pour tout  $a$  différent de 0 et  $\pm 1$  nous avons

$$[a, -a, a^2, 1] = \frac{(a - a^2)(-a - 1)}{(-a - a^2)(a - 1)} = -1. \quad (23.239) \quad \text{EQooUPTOooOsEXip}$$

Et en prenant  $\alpha(a)$  en guise de  $a$  nous avons aussi

$$[\alpha(a), -\alpha(a), \alpha(a)^2, \alpha(1)] = -1. \quad (23.240) \quad \text{EQooXYKYooQJaiMR}$$

Vu que  $\alpha$  préserve les divisions harmoniques, l'équation (23.239) donne aussi

$$[\alpha(a), \alpha(-a), \alpha(a^2), \alpha(1)] = -1, \quad (23.241)$$

c'est-à-dire

$$[\alpha(a), -\alpha(a), \alpha(a^2), 1] = -1. \quad (23.242) \quad \text{EQooYYHLooSELBp}$$

Comparant (23.240) avec (23.242) et en tenant compte de l'unicité du birapport<sup>18</sup> nous avons

$$\alpha(a^2) = \alpha(a)^2. \quad (23.243)$$

Avec cela nous pouvons y aller en remarquant que

$$ab = \left(\frac{a + b}{2}\right)^2 - \left(\frac{a - b}{2}\right)^2. \quad (23.244)$$

Nous appliquons  $\alpha$  à cette dernière équations en tenant compte de ce que nous savons déjà

$$\alpha(ab) = \left(\frac{\alpha(a) + \alpha(b)}{2}\right)^2 - \left(\frac{\alpha(a) - \alpha(b)}{2}\right)^2 = \alpha(a)\alpha(b). \quad (23.245)$$

□

17. Définition 1.96.

18. C'est-à-dire que si trois éléments du birapport sont donnés, le quatrième est fixé. C'est une variation sur la thème de la proposition 23.48(2).

**Théorème 23.95** ([680, 681]).

Le groupe circulaire de  $\mathbb{C}$  est le groupe des bijections  $\hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$  préservant l'ensemble des cercles-droites.

*Démonstration.* L'inclusion dans un sens est facile : les homographies conservent l'ensemble des cercles et droites par la proposition 23.80. Et la conjugaison complexe aussi.

Soit une bijection  $f: \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$  préservant les cercles-droites. Nous supposons dans un premier temps que  $f(0) = 0$ ,  $f(1) = 1$  et  $f(\infty) = \infty$ .

(1) **Pour  $f$  vérifiant  $f(0, 1, \infty) = 0, 1, \infty$** 

Si  $\mathcal{C}$  est un cercle alors  $f(\mathcal{C})$  est un cercle ou une droite, mais vu que  $\mathcal{C}$  ne contient pas  $\infty$ , l'ensemble  $f(\mathcal{C})$  ne le contient pas non plus. Donc  $f$  transforme un cercle en un cercle et une droite en une droite.

(1a)  **$f$  préserve les divisions harmoniques**

Soient  $a, b, c, x$  dans  $\hat{\mathbb{C}}$  tels que  $[a, b, c, x] = -1$ . Nous allons prouver que  $[f(a), f(b), f(c), f(x)] = -1$ .

Si  $a, b$  et  $c$  sont colinéaires, nous suivons la construction du lemme 23.92. Soit  $m$  hors de la droite  $(ab)$  et une droite  $D$  passant par  $c$  et coupant  $[ma]$  en  $p$  et  $[mb]$  en  $q$ . Nous posons  $n = (pb) \cap (qa)$ . Alors  $x = (mn) \cap (ac)$ .

L'application  $f$  est une bijection qui respecte les intersections, les tangences, les cercles et les droites). Le point  $f(m)$  est hors de la droite  $(f(a)f(b))$ . La droite  $f(D)$  passe par  $f(c)$  et coupe les segments  $[f(m)f(a)]$  en  $f(p)$  et  $[f(m)f(b)]$  en  $f(q)$ . Alors nous avons

$$f(n) = (f(p)f(b)) \cap (f(q)f(a)) \quad (23.246)$$

et aussi

$$f(x) = (f(m)f(n)) \cap (f(a)f(c)) \quad (23.247)$$

Donc  $f(x)$  se construit à partir de  $f(a)$ ,  $f(b)$  et  $f(c)$  en suivant la même construction que  $x$  à partir de  $a$ ,  $b$  et  $c$ . Nous en concluons que  $[f(a), f(b), f(c), f(x)] = -1$ .

Si  $a, b$  et  $c$  sont cocycliques, le même raisonnement, en suivant le lemme 23.91 nous donne le même résultat.

(1b)  **$f$  est un automorphisme du corps  $\mathbb{C}$** 

C'est le lemme 23.94.

(1c) **Et enfin ...**

Notre application  $f$  est un automorphisme du corps  $\mathbb{C}$  qui fixe  $\mathbb{R}$  parce qu'elle laisse invariante les droites dans  $\mathbb{C}$ . Donc la proposition 6.6 nous dit que  $f$  est soit l'identité soit la conjugaison complexe. Dans les deux cas,  $f$  est dans le groupe circulaire.

(2) **Pour  $f$  plus générale**

Nous ne supposons plus que  $f$  fixe 0, 1 et  $\infty$ . En tout cas les nombres  $f^{-1}(1)$ ,  $f^{-1}(0)$  et  $f^{-1}(\infty)$  sont distincts parce que  $f$  est une bijection. Nous pouvons considérer une homographie<sup>19</sup>  $\phi: \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$  telle que  $\phi(1) = f^{-1}(1)$ ,  $\phi(0) = f^{-1}(0)$  et  $\phi(\infty) = f^{-1}(\infty)$ . Dans ce cas l'application

$$g = f \circ \phi \quad (23.248)$$

vérifie  $g(1) = 1$ ,  $g(0) = 0$  et  $g(\infty) = \infty$  tout en continuant à transformer un cercle-droite en un cercle-droite. Donc  $f \circ \phi$  est soit l'identité soit la conjugaison complexe. Avec ça, l'application

$$f = g \circ \phi^{-1} \quad (23.249)$$

est la composée d'une homographie avec soit l'identité soit la conjugaison complexe. Elle est donc dans le groupe circulaire.

□

19. Attention : ici nous parlons d'homographies de  $\hat{\mathbb{C}}$ , pas de  $P(\mathbb{C}^2)$ . L'existence d'une telle application demande de composer le corolaire 23.35 avec l'application  $\varphi_0$  et la définition 23.74 et 23.75.

### 23.6.6 Action du groupe modulaire

Le **demi-plan de Poincaré** est l'ensemble

$$P = \{z \in \mathbb{C} \text{ tel que } \Im(z) > 0\}. \quad (23.250)$$

Le **groupe modulaire** est le quotient de groupes

$$\mathrm{PSL}(2, \mathbb{Z}) = \frac{\mathrm{SL}(2, \mathbb{Z})}{\mathbb{Z}_2}. \quad (23.251)$$

Ce sont donc les matrices au signe près de la forme

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (23.252)$$

où  $a, b, c$  et  $d$  sont entiers tels que  $ad - cb = 1$ .

**Théorème 23.96** ([682]).

Le groupe modulaire agit fidèlement (définition 2.31) sur le demi-plan de Poincaré par

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} * z = \frac{az + b}{cz + d}. \quad (23.253)$$

L'ensemble  $D = D_1 \cup D_2$  avec

$$D_1 = \{z \in P \text{ tel que } |z| > 1, -\frac{1}{2} \leq \Re(z) < \frac{1}{2}\} \quad (23.254a)$$

$$D_2 = \{z \in P \text{ tel que } |z| = 1, -\frac{1}{2} \leq \Re(z) \leq 0\} \quad (23.254b)$$

est un domaine fondamental (définition 2.38) de cette action.

De plus si nous notons

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad (23.255)$$

alors pour tout  $z \in P$ , il existe  $A \in \mathrm{gr}(S, T)$  telle que  $A * z \in D$ .

*Démonstration.* Nous divisons la preuve en plusieurs étapes.

(1) **Bien définie**

D'abord il faut remarquer que l'action (23.253) est bien définie par rapport au quotient :  $A * z = (-A) * z$ . La vérification est immédiate.

(2) **Interne**

Montrons que si  $A \in \mathrm{PSL}(2, \mathbb{Z})$  et  $z \in P$  alors  $A * z \in P$ . Nous avons

$$A * z = \frac{az + b}{cz + d} = \frac{(az + b)(c\bar{z} + d)}{|cz + d|^2} = \frac{a|z|c + azd + bc\bar{z} + bd}{|cz + d|^2}, \quad (23.256)$$

et donc en décomposant  $z = \Re(z) + i\Im(z)$ ,

$$\Im(A * z) = \Im\left(\frac{azd + bc\bar{z}}{|cz + d|^2}\right) = \frac{ad - bc}{|cz + d|^2} \Im(z) = \frac{\Im(z)}{|cz + d|^2} \quad (23.257)$$

où nous avons tenu compte de  $ad - bc = 1$ . Donc l'action respecte la (stricte) positivité de la partie imaginaire.

(3) **Action**

Nous vérifions maintenant que la formule donne bien une action :  $A * (B * z) = (AB) * z$ . Cela est un bon calcul :

$$A * (B * z) = A * \left( \frac{a'z + b'}{c'z + d'} \right) \quad (23.258a)$$

$$= \frac{a \left( \frac{a'z + b'}{c'z + d'} \right) + b}{c \left( \frac{a'z + b'}{c'z + d'} \right) + d} \quad (23.258b)$$

$$= \frac{a(a'z + b') + b(c'z + d')}{c(a'z + b') + d(c'z + d')} \quad (23.258c)$$

$$= \frac{(aa' + bc')z + (ab' + bd')}{(ca' + dc')z + (cb' + dd')} \quad (23.258d)$$

$$= \begin{pmatrix} aa' + bc' & ab' + bd' \\ a'c + dc' & cb' + dd' \end{pmatrix} * z \quad (23.258e)$$

$$= (AB) * z. \quad (23.258f)$$

(4) **Fidèle**

Soit  $A \in \text{PSL}(2, \mathbb{Z})$  tel que pour tout  $z \in P$  nous ayons

$$\frac{az + b}{cz + d} = z. \quad (23.259)$$

Alors nous avons

$$cz^2 + (d - a)z + b = 0. \quad (23.260)$$

Cela est donc un polynôme en  $z$  qui s'annule sur un ouvert<sup>20</sup> (le demi-plan de Poincaré). Il doit donc être identiquement nul, donc  $c = b = a - d = 0$ . Si vous n'y croyez pas, écrivez pour  $z = \epsilon i$  (avec  $\epsilon > 0$ ) :

$$-c\epsilon^2 + \epsilon(d - a)i + b = 0 \quad (23.261)$$

pour tout  $\epsilon$ . Le fait d'avoir  $c\epsilon^2 = b$  pour tout  $\epsilon$  implique que  $c = b = 0$ . Donc  $A$  est de la forme

$$A = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \quad (23.262)$$

avec la contrainte supplémentaire que  $ad = 1$ , les nombres  $a$  et  $b$  étant entiers. Nous avons donc soit  $a = d = 1$  soit  $a = d = -1$ . Étant donné le quotient par  $\mathbb{Z}_2$ , ces deux possibilités donnent le même élément de  $\text{PSL}(2, \mathbb{Z})$ .

(5) **Les orbites intersectent  $D$** 

Soit  $z \in P$ . Nous devons trouver  $A \in \text{PSL}(2, \mathbb{Z})$  tel que  $A * z \in D$ . Nous savons déjà que

$$\Im(A * z) = \frac{\Im(z)}{|cz + d|^2}. \quad (23.263)$$

Nous notons  $\mathcal{O}_z$  l'orbite de  $z$  sous le groupe modulaire et nous posons

$$I_z = \{\Im(u) \text{ tel que } u \in \mathcal{O}_z\} = \{\Im(A * z) \text{ tel que } A \in \text{PSL}(2, \mathbb{Z})\}, \quad (23.264)$$

l'ensemble des parties imaginaires des éléments de l'orbite de  $z$ . Nous allons montrer que cet ensemble est borné vers le haut en montrant que la quantité  $|cz + d|$  ne peut, à  $z$  donné, prendre qu'un nombre fini de valeurs plus grandes que  $\Im(z)$ <sup>21</sup>. Nous cherchons donc les couples  $(c, d) \in \mathbb{Z}^2$  tels que  $|cz + d| < 1$ .

20. On ne peut pas dire que  $b = 0$  simplement en justifiant qu'on l'obtient en posant  $z = 0$  parce que  $z = 0$  n'est pas dans le demi-plan de Poincaré.

21. Bien que cela ne soit pas indispensable pour la preuve, remarquons que  $I_z$  ne comprend qu'une quantité au plus dénombrable de valeurs. Le fait que, à  $z$  donné, la quantité  $|cz + d|^2$  puisse être rendue aussi grande que l'on veut est évident. Donc  $I_z$  est borné vers le bas par zéro (qui n'est pas atteint, mais qui est une valeur d'adhérence).



Nous avons  $\Im(cz + d) = c\Im(z)$ , donc  $|cz + d| \geq |c\Im(z)|$ , mais il n'y a qu'un nombre fini de  $c \in \mathbb{Z}$  tels que  $|c\Im(z)| < 1$ . De la même façon, pour la partie réelle nous avons

$$\Re(cz + d) = c\Re(z) + d, \quad (23.265)$$

et pour chaque  $c$ , il n'y a qu'un nombre fini de  $d \in \mathbb{Z}$  qui laissent cette quantité plus petite que 1 (en valeur absolue).

Donc  $I_z$  possède un maximum. Soit  $A_1 \in \text{PSL}(2, \mathbb{Z})$  tel que  $\Im(A_1 * z) = \max I_z$ . Nous notons  $z_1 = A_1 * z$ , et que nous n'avons à priori pas l'unicité. Nous allons maintenant agir sur  $z_1$  avec l'élément

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad (23.266)$$

pour ramener  $z_1$  dans le domaine  $D$ . Si  $u \in P$  nous avons  $T * u = u + 1$  et donc

$$T^n * u = u + n. \quad (23.267)$$

Vu que  $D$  est de largeur 1, il existe un  $n$  (éventuellement négatif) tel que

$$\Re(T^n * z_1) \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (23.268)$$

Notons qu'ici le fait d'être ouvert d'un côté et fermé de l'autre joue de façon essentielle (pour l'unicité aussi). Nous notons  $z_2 = T^n * z_1$ .

Supposons un instant que  $|z_2| < 1$ . Nous considérons l'élément

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (23.269)$$

qui fait

$$\Im(S * z) = \frac{\Im z}{|z|^2}. \quad (23.270)$$

Donc si  $|z_2| < 1$  alors  $\Im(S * z_2) > \Im(z_2)$ , ce qui contredit la maximalité de  $\Im(z_2)$  dans  $I_z$ . Nous en déduisons que  $|z_2| \geq 1$ .

Si  $|z_2| > 1$ , alors  $z_2 \in D_1$  et c'est bon. Si  $|z_2| = 1$ , alors il faut encore un peu travailler. Si  $z_2 \pm 1$  est à l'intérieur du disque, alors en agissant avec  $T$  ou  $T^{-1}$  nous retrouvons la même contradiction que précédemment. En écrivant  $z_2 = e^{i\theta}$ , nous devons donc avoir  $2 \cos(\theta) \leq 1$  ou encore  $|\Re(z_2)| \leq \frac{1}{2}$ . Donc si  $\Re(z_2) \leq 0$  alors  $z_2 \in D_2$ .

Le dernier cas à traiter est  $\Re(z_2) \in ]0, \frac{1}{2}]$ , c'est-à-dire  $\theta \in [\frac{\pi}{3}, \frac{\pi}{2}[$ . Dans ce cas l'action avec  $S$  ramène l'angle dans la bonne zone parce que  $S * z = -\frac{1}{z}$  et donc  $S * (\rho e^{-i\theta}) = -\frac{1}{\rho} e^{-i\theta}$ .

## (6) Unicité

Nous voulons à présent montrer que si  $z \in D$ , alors  $A * z$  n'est plus dans  $D$  (sauf si  $A = \pm 1$ ). Nous supposons que  $z \in D$  et  $A \in \text{PSL}(2, \mathbb{Z})$  soient tels que  $A * z \in D$ , et nous prouvons qu'alors soit nous arrivons à une contradiction soit nous arrivons à  $A = 1$ . Pour cela nous allons décomposer en de nombreux cas.

(6a) Nous commençons par  $\Im(A * z) \geq \Im(z)$ . Dans ce cas nous avons  $|cz + d| \leq 1$  et en particulier  $|c||\Im(z)| \leq 1$ . Étant donné que le point de  $D$  qui a la partie imaginaire la plus petite est  $-\frac{1}{2} + \frac{2}{\sqrt{3}}i$ , nous trouvons  $|c| \leq 2/\sqrt{3}$ . Vu que  $c$  doit être entier, nous avons trois cas :  $c = -1, 0, 1$ .

(6a.i) Soit  $c = 0$ . Alors  $A = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$  et la condition de déterminant est  $ad = 1$ , ce qui signifie  $a = d = 1$  (la possibilité  $a = b = -1$  est « éliminée » le quotient par  $\mathbb{Z}_2$  définissant  $\text{PSL}(2, \mathbb{Z})$ ). La matrice  $A$  doit alors être de la forme

$$A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \quad (23.271)$$

et  $A * z = z + b$ . Si  $z \in D$ , alors le seul  $z + b$  à être (peut-être) encore dans  $D$  est  $b = 0$ , mais alors  $A$  est l'identité.

(6a.ii) Soit  $c = 1$ . Alors la condition  $|cz + d| \leq 1$  nous donne trois possibilités<sup>22</sup> :  $d = -1, 0, 1$ .

- A. Si  $d = -1$ , alors nous devons avoir  $|z - 1| \leq 1$ . Il est instructif de faire un dessin, mais le point d'intersection entre les cercles  $|z| = 1$  et  $|z - 1| = 1$  est le point  $\frac{1}{2} + \frac{\sqrt{3}}{2}i$ , qui n'est pas dans  $D$ . Bref, il n'y a pas de points dans  $D$  vérifiant  $|z - 1| \leq 1$ .
- B. Si  $d = 1$ , alors (et c'est maintenant que la dissymétrie de  $D$  intervient) nous avons le point

$$z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \quad (23.272)$$

qui est dans  $D$  et qui vérifie  $|z + 1| \leq 1$ . Voyons à quoi ressemble la matrice  $A$  dans ce cas. Son déterminant est  $a - b = 1$ . Nous écrivons donc

$$A = \begin{pmatrix} b+1 & b \\ 1 & 1 \end{pmatrix}, \quad (23.273)$$

et en tenant compte du fait que  $z\bar{z} = |z + 1| = 1$ , nous calculons

$$A * z = \frac{(b+1)z + b}{z + 1} \quad (23.274a)$$

$$= \frac{(bz + z + b)(\bar{z} + 1)}{|z + 1|^2} \quad (23.274b)$$

$$= z + b + 1. \quad (23.274c)$$

La seule façon de ne pas quitter  $D$  est d'avoir  $b = -1$ , mais alors nous avons

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \quad (23.275)$$

et  $A * z = z$ . Donc au final  $z$  est quand même le seul de son orbite à être dans  $D$ .

Notons au passage cette très intéressante propriété du point

$$z_0 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i. \quad (23.276)$$

C'est un point de qui vérifie  $z_0 = A * z_0$  pour un élément non trivial  $A$  de  $PSL(2, \mathbb{Z})$ . L'existence d'un tel élément est ce qui va nous coûter un peu de sueur pour prouver que  $PSL(2, \mathbb{Z})$  est engendré par  $S$  et  $T$ .

- C. Le cas  $d = 0$  nous fait écrire  $1 = \det A = -b$ , donc  $b = -1$  et

$$A = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}. \quad (23.277)$$

Nous avons alors  $A * z = a - \frac{1}{z}$ . De plus la condition  $|z| \leq 1$  revient à  $|z| = 1$ . Pour les nombres complexes de module 1, l'opération  $z \rightarrow -1/z$  est la symétrie autour de l'axe des imaginaires purs. Le seul à ne pas sortir de  $D$  est le fameux  $z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ , qui revient sur lui-même avec  $a = -1$ .

Nous passons à la possibilité  $c = -1$ . Dans ce cas la matrice est de la forme

$$A = \begin{pmatrix} a & b \\ -1 & d \end{pmatrix}, \quad (23.278)$$

et nous revenons au cas  $c = 1$  en prenant  $-A$  au lieu de  $A$ .

---

22. Je ne rigolais pas quand je disais qu'on allait avoir de nombreux cas.

(6b) Nous passons au cas  $\Im(A * z) < \Im(z)$ . Nous récrivons cette condition avec

$$\Im(A * z) < \Im(A^{-1} * (A * z)). \quad (23.279)$$

Si nous supposons que  $z$  et  $A$  sont tels que  $z$  et  $A * z$  soient tous deux dans  $D$ , alors  $z' = A * z$  est un élément de  $D$  tel que

$$\Im(z') < \Im(A^{-1} * z'). \quad (23.280)$$

Or nous avons vu qu'aucun élément de  $D$  vérifiant cette condition n'existait sans être trivial (celui qui ne bouge pas). Pour cela il suffit d'appliquer tout ce que nous venons de dire avec  $A^{-1}$  au lieu de  $A$ .

### (7) Quelques conclusions

Après avoir passé tous les cas en revue, le fameux point  $z_0 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$  est l'unique point de  $D$  à accepter une matrice non triviale  $A \in \text{PSL}(2, \mathbb{Z})$  telle que  $z_0 = A * z_0$ .

Nous remarquons aussi que tous les points de  $P$  sont ramenés dans  $D$  par une matrice obtenue comme produit de  $T$ ,  $S$ ,  $T^{-1}$  et  $S^{-1}$ .

□

CorJQwgNp

### **Corolaire 23.97** ([683]).

Les matrices  $S$  et  $T$  génèrent le groupe modulaire au sens où toute matrice de  $\text{PSL}(2, \mathbb{Z})$  s'écrit comme

$$T^{m_1} S^{p_1} \dots T^{m_k} S^{p_k} \quad (23.281)$$

pour un certain  $k$  et des nombres  $m_i, p_i \in \mathbb{Z}$ . Autrement dit,  $\text{PSL}(2, \mathbb{Z}) = \text{gr}(S, T)$ .

*Démonstration.* Soit  $z$ , un point de  $D$  autre que  $z_0$ . Alors si  $A \in \text{PSL}(2, \mathbb{Z})$  est non trivial nous avons  $A * z$  hors de  $D$ . Du coup, comme vu dans la démonstration du théorème 23.96, il existe  $B \in \text{gr}(S, T)$  tel que  $B * (A * z) \in D$ . Vu que  $D$  ne contient qu'un seul point de chaque orbite, nous avons

$$B * A * z = z, \quad (23.282)$$

et donc  $BA = \pm \mathbb{1}$ , ce qui prouve que <sup>23</sup>  $A = B^{-1}$ , c'est-à-dire que  $A \in \text{gr}(S, T)$ . □

---

23. Dans  $\text{PSL}(2, \mathbb{Z})$ , nous n'avons pas besoin de mettre  $\pm$  parce qu'il est compris dans la définition.

# Chapitre 24

## Analyse vectorielle

### 24.1 Le théorème de Green

Soit un champ de vecteurs

$$F(x, y, z) = \begin{pmatrix} F_1(x, y, z) \\ F_2(x, y, z) \\ F_3(x, y, z) \end{pmatrix} \quad (24.1)$$

et un chemin  $\sigma: [a, b] \rightarrow \mathbb{R}^3$  donné par

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}. \quad (24.2)$$

Nous avons défini la circulation de  $F$  le long de  $\sigma$  par

$$\begin{aligned} \int_{\sigma} F \cdot d\sigma &= \int_a^b F(\sigma(t)) \cdot \sigma'(t) dt \\ &= \int_a^b \left[ F_1(\sigma(t))x'(t) + F_2(\sigma(t))y'(t) + F_3(\sigma(t))z'(t) \right] dt \\ &= \int_{\sigma} F_1 dx + F_2 dy + F_3 dz. \end{aligned} \quad (24.3)$$

La dernière ligne est juste une notation compacte<sup>1</sup>. Elle sert à se souvenir qu'on va mettre  $x'$  à côté de  $F_1$ ,  $y'$  à côté de  $F_2$  et  $z'$  à côté de  $F_3$ . L'avantage de cette notation est qu'on peut écrire d'autres combinaisons.

Si  $f$  et  $g$  sont deux fonctions sur  $\mathbb{R}^3$ , nous pouvons écrire

$$\int_{\sigma} f dy + g dz. \quad (24.4)$$

Cela signifie

$$\int_a^b \left[ f(\sigma(t))y'(t) + g(\sigma(t))z'(t) \right] dt. \quad (24.5)$$

Soit  $D$  une région du plan et  $\sigma$ , son contour que nous prenons, par convention<sup>2</sup>, dans l'orientation trigonométrique, comme indiqué sur la figure 24.1. Nous supposons également que le domaine  $D$  n'a pas de trous intérieurs.

Nous notons par  $\sigma = \partial D$  le bord de  $D$ , c'est-à-dire le contour dont nous venons de parler.

---

1. Il y aurait beaucoup de choses à dire là-dessus, mais la vie est trop courte pour parler de formes différentielles, et c'est dommage.

2. Il y aurait beaucoup de choses à dire sur ça aussi, mais...

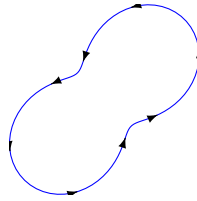


FIGURE 24.1: Un contour avec son orientation.

LabelFigVDFMooHMmFZr

**Théorème 24.1** (Théorème de Green).*Soient  $P, Q: D \rightarrow \mathbb{R}$  deux fonctions de classe  $C^1$ . Alors*

$$\int_{\partial D} P dx + Q dy = \int_D \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy. \quad (24.6)$$

EqThGreen

Pour rappel, l'intégrale du membre de gauche signifie

$$\int_a^b \left[ P(\sigma(t)) \sigma'_x(t) + Q(\sigma(t)) \sigma'_y(t) \right] dt. \quad (24.7)$$

Ce n'est d'ailleurs rien d'autre que l'intégrale du champ de vecteurs  $\begin{pmatrix} P \\ Q \end{pmatrix}$ .**Corolaire 24.2.***L'aire du domaine  $D$  est donnée par*

$$A = \frac{1}{2} \int_{\partial D} (x dy - y dx). \quad (24.8)$$

*Démonstration.* L'intégrale  $\int_{\partial D} (x dy - y dx)$  se traite avec le théorème de Green où l'on pose  $P = -y$  et  $Q = x$ . Nous avons donc

$$\begin{aligned} \int_{\partial D} -y dx + x dy &= \int_D \left( \frac{\partial x}{\partial x} - \frac{\partial(-y)}{\partial y} \right) dx dy \\ &= \int_D 2 dx dy. \end{aligned} \quad (24.9)$$

La dernière ligne est bien le double de la surface. □**Exemple 24.3.**Calculons (encore une fois) l'aire du disque de rayon  $R$ . Il s'agit de calculer l'intégrale

$$I = \frac{1}{2} \int_{\sigma} (x dt - y dx) \quad (24.10)$$

où  $\sigma$  est le cercle donné par

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} R \cos(t) \\ R \sin(t) \end{pmatrix} \quad (24.11)$$

Le calcul est

$$\begin{aligned} I &= \frac{1}{2} \int_0^{2\pi} \underbrace{R \cos(\theta)}_x \underbrace{R \cos(\theta)}_{y'} - \underbrace{R \sin(\theta)}_y \underbrace{(-R \sin(\theta))}_{x'} d\theta \\ &= \frac{R^2}{2} \int_0^{2\pi} d\theta \\ &= \pi R^2. \end{aligned} \quad (24.12)$$

△

**Exemple 24.4.**

Calculons l'aire de l'ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \quad (24.13)$$

dont le bord est donné par

$$\begin{cases} x(t) = a \cos(t) \\ y(t) = b \sin(t). \end{cases} \quad (24.14a)$$

$$(24.14b)$$

Le terme  $xdy$  devient  $a \cos(t)b \cos(t) = ab \cos^2(t)$  et le terme  $ydx$  devient  $b \sin(t)(-a \sin(t)) = -ab \sin^2(t)$ . L'intégrale qui donne la surface est donc

$$\frac{1}{2} \int_{\partial D} (xdy - ydx) = \frac{1}{2} \int_0^{2\pi} ab = \pi ab. \quad (24.15)$$

△

Le théorème de Green peut être mis sous une autre forme.

ThoGreenVecto

**Théorème 24.5** (Théorème de Green, forme vectorielle).

Si  $G$  est un champ de vecteurs sur  $D$ , nous avons

$$\int_{\partial D} G \cdot d\sigma = \int_D (\nabla \times G) \cdot dS \quad (24.16)$$

EqGreenVecto

où le second membre est le flux de  $\nabla \times G$  sur la surface  $D$ .

*Démonstration.* Analysons le membre de droite. Nous savons que  $D$  est une surface dans le plan  $\mathbb{R}^2$ . Le vecteur normal à la surface est donc simplement le vecteur (constant)  $e_z$ . Le produit scalaire  $(\nabla \times F) \cdot dS$  est donc  $(\nabla \times F) \cdot e_z$  et se réduit à la troisième composante du rotationnel, c'est-à-dire

$$\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}. \quad (24.17)$$

Cela est bien le membre de droite de l'équation (24.6). Le membre de gauche de cette dernière est bien le membre de gauche de (24.16). □

ExempleGreenSql

**Exemple 24.6.**

Soit le champ de vecteurs  $F(x, y) = \begin{pmatrix} xy^2 \\ y + x \end{pmatrix}$ , et soit à calculer

$$\int_D \nabla \times F \cdot dS \quad (24.18)$$

où  $D$  est la région comprise entre les courbes  $y = x^2$  et  $y = x$  pour  $x \geq 0$  (voir la figure 24.2).

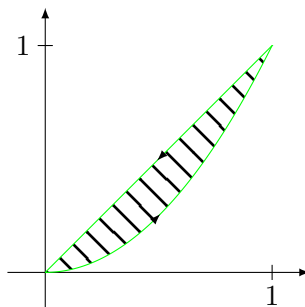


FIGURE 24.2: Le contour d'intégration pour l'exemple 24.6. LabelFigLLVMooWOkvAB

Nous pouvons calculer cette intégrale directement en calculant le rotationnel de  $F$  :

$$\nabla \times F = \begin{pmatrix} 0 \\ 0 \\ 1 - 2xy \end{pmatrix}. \quad (24.19)$$

Par conséquent l'intégrale à effectuer est

$$I = \int_0^1 dx \int_{x^2}^x (1 - 2xy) dy = \frac{1}{12}. \quad (24.20)$$

```
-----
| Sage Version 4.6.1, Release Date: 2011-01-11 |
| Type notebook() for the GUI, and license() for information. |
|-----
```

```
sage: f(x,y)=1-2*x*y
sage: f.integrate(y,x**2,x).integrate(x,0,1)
(x, y) |--> 1/12
```

L'autre façon de calculer l'intégrale est d'utiliser le théorème de Green et de calculer la circulation de  $F$  le long de  $\partial D$  :

$$I = \int_{\partial D} F \cdot \sigma. \quad (24.21)$$

Le chemin  $\sigma = \partial D$  est composé de la parabole  $y = x^2$  et du segment de droite  $x = y$ . Attention : il faut respecter l'orientation. Nous avons

$$\sigma_1(t) = (t, t^2) \quad (24.22)$$

et

$$\sigma_2(t) = (1 - t, 1 - t). \quad (24.23)$$

Notez bien que le second chemin est  $(1 - t, 1 - t)$  et non  $(t, t)$  parce qu'il faut le parcourir dans le bon sens (voir le dessin).

Commençons par le premier chemin :

$$\begin{aligned} \sigma_1(t) &= (t, t^2) \\ \sigma_1'(t) &= (1, 2t) \\ F(\sigma_1(t)) &= \begin{pmatrix} t^5 \\ t + t^2 \end{pmatrix}, \end{aligned} \quad (24.24)$$

et par conséquent

$$F(\sigma_1(t)) \cdot \sigma_1'(t) = t^5 + 2t^2 + 2t^3, \quad (24.25)$$

et le premier morceau de la circulation vaut

$$\int_{\sigma_1} F \cdot d\sigma_1 = \int_0^1 t^5 + 2t^2 + 2t^3 = \frac{4}{3}. \quad (24.26)$$

Pour le second chemin :

$$\begin{aligned} \sigma_2(t) &= (1 - t, 1 - t) \\ \sigma_2'(t) &= (-1, -1) \\ F(\sigma_2(t)) &= \begin{pmatrix} (1 - t)^3 \\ 2(1 - t) \end{pmatrix}. \end{aligned} \quad (24.27)$$

Par conséquent

$$F(\sigma_2(t)) \cdot \sigma_2'(t) = -(1 - t)^2 - 2(1 - t). \quad (24.28)$$

Le second morceau de la circulation est par conséquent

$$\int_0^1 -(1-t)^2 - 2(1-t)dt = -\frac{5}{4}. \quad (24.29)$$

La circulation de  $F$  le long de  $\sigma$  est donc égale à

$$\frac{4}{3} - \frac{5}{4} = \frac{1}{12}. \quad (24.30)$$

Comme prévu, nous obtenons le même résultat.  $\triangle$

## 24.2 Théorème de la divergence dans le plan

### 24.2.1 La convention de sens de parcours

Soient  $D$ , un domaine dans le plan et un paramétrage

$$\begin{aligned} \sigma: [a, b] &\rightarrow \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \end{aligned} \quad (24.31)$$

un paramétrage du bord  $\partial D$  de  $D$ . La normale à  $\sigma$  est perpendiculaire à la tangente, donc la normale extérieure de norme 1 vaut

$$n = \frac{(y'(t), -x'(t))}{\sqrt{(x'(t))^2 + (y'(t))^2}} \quad \text{ou} \quad n = -\frac{(y'(t), -x'(t))}{\sqrt{(x'(t))^2 + (y'(t))^2}}. \quad (24.32)$$

Comment faire le choix ?

Nous prenons comme convention que le sens *du chemin* doit être tel que le vecteur normal extérieur soit

$$n = \frac{(y'(t), -x'(t))}{\sqrt{(x'(t))^2 + (y'(t))^2}}. \quad (24.33)$$

Donc si le chemin  $\sigma$  donne lieu à un vecteur  $n$  pointant vers l'intérieur, il faut utiliser le chemin qui va dans le sens contraire :  $\tilde{\sigma}(t) = \sigma(1-t)$ .

Les vecteurs tangents et normaux d'un contour sont dessinés sur la figure 24.3.

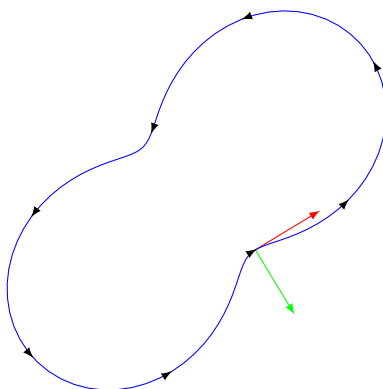


FIGURE 24.3: Le champ de vecteurs tangents est dessiné en rouge tandis qu'en vert nous avons le champ de vecteurs normaux extérieurs.

LabelFigDDCTooYscVzA



### 24.2.2 Théorème de la divergence

**Théorème 24.7** (Théorème de la divergence).

Soit  $F$  un champ de vecteurs sur  $\mathbb{R}^2$ . Le flux de  $F$  à travers le bord de  $D$  est égal à l'intégrale de la divergence de  $F$  sur  $D$ . En formule :

$$\int_{\partial D} F \cdot n \, d\sigma = \int_D \nabla \cdot F \, dx dy. \quad (24.34)$$

*Démonstration.* Tant  $F \cdot n$  que  $\nabla \cdot F$  sont des fonctions. Le membre de gauche est donc l'intégrale d'une fonction sur un chemin et le membre de droite est l'intégrale d'une fonction sur une surface. Notre convention de sens de parcours du chemin permet d'écrire le produit scalaire  $F \cdot n$  sous la forme suivante :

$$\begin{aligned} F \cdot n &= \frac{1}{\|\sigma'\|} \begin{pmatrix} F_x \\ F_y \end{pmatrix} \cdot \begin{pmatrix} y' \\ -x' \end{pmatrix} \\ &= \frac{1}{\|\sigma'\|} (F_x y' - F_y x') \\ &= \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \\ &= \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'. \end{aligned} \quad (24.35)$$

Par conséquent, la fonction

$$F \cdot n \quad (24.36)$$

est la même que la fonction

$$\frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'. \quad (24.37)$$

L'intégrale de cette dernière fonction sur le chemin  $\sigma$  est

$$\begin{aligned} I &= \int_{\sigma} F \cdot n \\ &= \int_{\sigma} \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma' \\ &= \int_a^b \frac{1}{\|\sigma'(t)\|} \begin{pmatrix} -F_y(\sigma(t)) \\ F_x(\sigma(t)) \end{pmatrix} \cdot \sigma'(t) \|\sigma'(t)\| dt \\ &= \int_a^b \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'(t) dt. \end{aligned} \quad (24.38)$$

Cette dernière intégrale est la circulation du champ de vecteurs  $\begin{pmatrix} -F_y \\ F_x \end{pmatrix}$  sur le chemin  $\sigma$ . Le théorème de Green 24.5 nous enseigne que la circulation le long d'un chemin est égale au flux du rotationnel à travers la surface. Par conséquent,

$$I = \int_D \left( \nabla \times \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \right) \cdot dS = \int_D \nabla \cdot F \, dx dy \quad (24.39)$$

□

## 24.3 Théorème de Stokes

Nous nous mettons maintenant dans  $\mathbb{R}^3$ , et nous y considérons une surface paramétrée  $S$  donc le bord est  $\partial S$ .

**Théorème 24.8** (Théorème de Stokes).

Alors le flux du rotationnel de  $F$  à travers  $S$  est égal à la circulation de  $F$  le long du bord. En formule :

$$\int_S \nabla \times F \cdot dS = \int_{\partial S} F \cdot d\sigma. \quad (24.40)$$

Nous pouvons nous donner une idée du pourquoi ce théorème est vrai. D'abord, si la surface est plate, cela est exactement le théorème de Green 24.5. Supposons maintenant que le bord reste plat, mais que la surface se déforme un petit peu. Le chemin

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 0 \end{pmatrix} \quad (24.41)$$

est tout autant le bord du disque plat de rayon 1 que celui de la demi-sphère

$$\phi(x, y) = \begin{pmatrix} x \\ y \\ \sqrt{1 - x^2 - y^2} \end{pmatrix}. \quad (24.42)$$

Le champ de vecteur que nous considérons est  $G = \nabla \times F$ . Il a un certain flux à travers le disque plat, et ce plus est égal à la circulation de  $F$  sur  $\sigma$ . Quel est le flux de  $G$  à travers la demi-sphère ? Étant donné que  $\nabla \cdot G = \nabla \cdot (\nabla \times F) = 0$ , le champ de vecteurs  $G$  est incompressible, de telle façon que tout ce qui rentre dans la demi-sphère doit en sortir. Le flux de  $G$  à travers la demi-sphère doit par conséquent être égal à celui à travers le disque plat.

**Exemple 24.9.**

Soit  $C$  l'intersection entre le cylindre  $x^2 + y^2 = 1$  et le plan  $x + y + z = 1$ . Calculer la circulation de

$$F(x, y, z) = \begin{pmatrix} -y^3 \\ x^3 \\ -z^3 \end{pmatrix} \quad (24.43)$$

le long de  $C$ .

Au lieu de calculer directement

$$\int_C F \cdot d\sigma, \quad (24.44)$$

nous allons calculer

$$\int_S \nabla \times F \cdot dS \quad (24.45)$$

où  $S$  est une surface dont  $C$  est le bord. Cette intégrale est à calculer avec la formule (20.294).

La première chose à faire est de trouver une surface dont le bord est  $C$  et en trouver un paramétrage  $\phi$ . Le plus simple est de prendre le graphe du plan sur le cercle  $x^2 + y^2 = 1$ . Un paramétrage de cette surface est simplement

$$\begin{aligned} \phi: D &\rightarrow \mathbb{R}^3 \\ (x, y) &\mapsto \begin{pmatrix} x \\ y \\ 1 - x - y \end{pmatrix} \end{aligned} \quad (24.46)$$

où  $D$  est le disque de rayon 1. Étant donné que cela paramètre le plan  $x + y + z - 1 = 0$ , le vecteur normal est  $n = e_x + e_y + e_z$ . Nous pouvons cependant calculer ce vecteur normal en suivant la recette usuelle. D'abord les vecteurs tangents sont

$$\frac{\partial \phi}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \frac{\partial \phi}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \quad (24.47)$$

Et le vecteur normal est donné par le produit vectoriel :

$$\begin{aligned} n &= \frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} \\ &= \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{vmatrix} \\ &= e_x + e_y + e_z. \end{aligned} \quad (24.48)$$

Ensuite, le rotationnel de  $F$  est donné par

$$\nabla \times F = 3(x^2 + y^2)e_z. \quad (24.49)$$

Par conséquent,

$$\nabla \times F \cdot \left( \frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} \right) = 3(x^2 + y^2). \quad (24.50)$$

L'intégrale à calculer est donc

$$\begin{aligned} \int_S \nabla \times F \cdot dS &= \int_D (\nabla \times F)(\phi(x, y)) \cdot \left( \frac{\partial \phi}{\partial x} \times \frac{\partial \phi}{\partial y} \right) dx dy \\ &= 3 \int_D (x^2 + y^2) dx dy. \end{aligned} \quad (24.51)$$

Cette dernière intégrale est l'intégrale d'une fonction sur le disque de rayon 1. Elle s'effectue en passant aux coordonnées polaires :

$$3 \int_D (x^2 + y^2) dx dy = 3 \int_0^{2\pi} d\theta \int_0^1 (r^2) r dr = \frac{3\pi}{2}. \quad (24.52)$$

△

## 24.4 Théorème de Gauss

Soit  $V$  une partie de  $\mathbb{R}^3$  délimitée par une surface  $S$  sur laquelle nous considérons la normale extérieure. Soit  $F$  un champ de vecteurs sur  $\mathbb{R}^3$ .

**Théorème 24.10** (Théorème de la divergence ou de Gauss).

*Le flux d'un champ de vecteur  $F$  à travers une surface fermée est égal à l'intégrale de la divergence sur le volume correspondant :*

$$\int_{\partial V} F \cdot dS = \int_V \nabla \cdot F dx dy dz. \quad (24.53)$$

Ce théorème signifie que la quantité de fluide qui s'accumule dans le volume (le flux est ce qui rentre moins ce qui sort) est égale à l'intégrale de  $\nabla \cdot F$  sur le volume, alors que nous savons que, localement, la quantité  $\nabla \cdot F(x, y, z)$  est la quantité de fluide qui s'accumule au point  $(x, y, z)$ .

**Remarque 24.11.**

Ce théorème ne fonctionne qu'avec des surfaces fermées. Essayer de l'appliquer au calcul de flux à travers des surfaces ouvertes n'a pas de sens parce qu'une surface ouverte ne délimite pas un volume.

**24.12.**

La formule de la divergence peut être utilisée comme intégration par partie. Si  $u$  est une fonction et  $F$  un champ de vecteurs,  $\nabla(uF) = \nabla(u) \cdot F + u\nabla \cdot F$  et alors

$$\int_{\partial V} uF \cdot n = \int_V \nabla(uF) = \int_V u\nabla \cdot F + \int_V F \cdot \nabla u \quad (24.54)$$

où  $n$  est le champ de vecteurs normal extérieur à  $V$ . En remettant les termes dans un ordre qui ressemble plus à l'intégration par partie :

$$\int_V F \cdot \nabla u = \int_{\partial V} u F \cdot n - \int_V u \nabla F. \quad (24.55)$$

**Exemple 24.13.**

Calculer le flux du champ de vecteurs

$$F(x, y, z) = \begin{pmatrix} 2x \\ y^2 \\ z^2 \end{pmatrix} \quad (24.56)$$

à travers la sphère de rayon 1 centrée à l'origine. Nous utilisons le théorème de la divergence

$$\int_S F \cdot n \, dS = \int_B \nabla \cdot F \, dx dy dz \quad (24.57)$$

où  $S$  est la sphère et  $B$  est la boule (la sphère pleine). La divergence de  $F$  se calcule :

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = 2 + 2x + 2y. \quad (24.58)$$

L'intégrale est donc en trois termes :

$$\begin{aligned} \int_B 2 &= 2 \text{Volume}(B) = \frac{8\pi}{3} \\ \int_B y \, dx dy dz &= 0 \\ \int_B z \, dx dy dz &= 0. \end{aligned} \quad (24.59)$$

△

Dans certains cas le théorème de Gauss permet de simplifier le calcul de l'intégrale d'une fonction sur une surface.

**Exemple 24.14.**

Soit à calculer l'intégrale

$$I = \int_{\partial B} (x^2 + y + z) dS, \quad (24.60)$$

c'est-à-dire l'intégrale de la fonction  $x^2 + y + z$  sur la sphère. Le vecteur normal à la sphère est

$$n = x e_x + y e_y + z e_z. \quad (24.61)$$

Étant donné que nous sommes sur la sphère de rayon 1, ce vecteur est même normé. La fonction que nous regardons n'est rien d'autre que  $F \cdot n$  avec

$$F = \begin{pmatrix} x \\ 1 \\ 1 \end{pmatrix}. \quad (24.62)$$

Nous pouvons donc simplement intégrer  $\nabla \cdot F$  sur toute la boule :

$$I = \int_B \nabla \cdot F \, dx dy dz = \int_B 1 \, dx dy dz = \frac{4\pi}{3}. \quad (24.63)$$

△

## 24.5 Coordonnées curvilignes

### 24.5.1 Base locale

Nous connaissons déjà les coordonnées sphériques et cylindriques sur  $\mathbb{R}^3$ . Ce sont des systèmes « un peu courbes ». Il en existe bien entendu de nombreux autres (sinon ce ne serait pas drôle), et nous allons faire une étude un peu générale de ces systèmes de **coordonnées curvilignes**. Des coordonnées curvilignes sur  $\mathbb{R}^3$  est n'importe quel <sup>3</sup> système qui permet de repérer un point de  $\mathbb{R}^3$  à partir de trois nombres.

Il s'agit donc d'un ensemble de trois applications

$$x_i: \mathbb{R}^3 \rightarrow \mathbb{R}. \quad (24.64)$$

Les coordonnées cylindriques sont

$$\begin{cases} x_1(r, \theta, z) = r \cos \theta & (24.65a) \\ x_2(r, \theta, z) = r \sin \theta & (24.65b) \\ x_3(r, \theta, z) = z & (24.65c) \end{cases}$$

Soit donc un système général  $q = (q_1, q_2, q_3)$  et

$$M(q) = \begin{pmatrix} x_1(q) \\ x_2(q) \\ x_3(q) \end{pmatrix}. \quad (24.66)$$

Si nous fixons  $q_2$  et  $q_3$  et que nous laissons varier  $q_1$ , nous obtenons une courbe <sup>4</sup> dont nous pouvons considérer le vecteur vitesse, c'est-à-dire le vecteur tangent. En chaque point nous avons ainsi trois vecteurs

$$\frac{\partial M}{\partial q_i}(q). \quad (24.67)$$

Nous disons que le système de coordonnées curviligne est **orthogonal** si ces trois vecteurs sont orthogonaux. Dans la suite nous supposons que c'est toujours le cas.

Nous posons

$$h_i = \left\| \frac{\partial M}{\partial q_i} \right\| \quad (24.68)$$

et nous considérons les trois vecteurs normés

$$e_i = h_i^{-1} \frac{\partial M}{\partial q_i}. \quad (24.69)$$

Les trois vecteurs  $\{e_1, e_2, e_3\}$  forment une base orthonormée dite **base locale**. Ce sont des vecteurs liés <sup>5</sup> au point  $M$ .

### 24.5.2 Importance de l'orthogonalité

Nous avons dit que nous nous restreignons au cas où les vecteurs  $e_i$  sont orthogonaux. En termes de produits scalaires, cela signifie

$$e_i \cdot e_j = \delta_{ij}. \quad (24.70)$$

Nous en étudions maintenant quelques conséquences. L'équation (24.69) peut s'écrire plus explicitement sous la forme

$$e_i = \sum_k h_i^{-1} \frac{\partial x_k}{\partial q_i} 1_k. \quad (24.71)$$

3. Nous n'entrons pas dans les détails de régularité.

4. Dans le cas des sphériques, c'est une demi-droite horizontale d'angle  $q_2$  et de hauteur  $q_3$ .

5. En géométrie différentielle on dira que ce sont des éléments de l'espace tangent, mais c'est une toute autre histoire.

Notez que pour chaque  $k$  et  $i$ , la quantité  $h_i^{-1} \frac{\partial x_k}{\partial q_i}$  est un simple nombre. Nous allons les mettre dans une matrice :

$$A_{ki} = h_i^{-1} \frac{\partial x_k}{\partial q_i}. \quad (24.72)$$

Cela nous donne le changement de base

$$e_i = \sum_k A_{ki} 1_k. \quad (24.73) \quad \text{EqChmBasisAkiAk}$$

Le produit  $e_i \cdot e_j$  s'écrit alors

$$\begin{aligned} e_i \cdot e_j &= \sum_{kl} A_{ki} A_{lj} \underbrace{1_k \cdot 1_l}_{=\delta_{kl}} \\ &= \sum_{kl} A_{ki} A_{lj} \delta_{kl} \\ &= \sum_k A_{ki} A_{kj} \\ &= \sum_k (A^T)_{ik} A_{kj}. \end{aligned} \quad (24.74)$$

Or cela doit valoir  $\delta_{ij}$ . Par conséquent

$$A^T = A^{-1}. \quad (24.75)$$

Le fait que les coordonnées curvilignes considérées soient orthogonales s'exprime donc par la fait que la matrice de changement de base est une matrice orthogonale.

Cette circonstance nous permet d'inverser le changement de base (24.73) en multipliant cette équation par  $(A^{-1})_{il}$  des deux côtés et en faisant la somme sur  $i$  :

$$\sum_i (A^{-1})_{il} e_i = \sum_{kl} A_{ki} \underbrace{(A^{-1})_{il}}_{=\delta_{kl}} 1_k, \quad (24.76)$$

par conséquent

$$\sum_i (A^T)_{il} e_i = 1_l, \quad (24.77)$$

et

$$1_l = \sum_i A_{li} e_i = \sum_i h_i^{-1} \frac{\partial x_l}{\partial q_i} e_i. \quad (24.78) \quad \text{EqChamvarunlAei}$$

Armés de cette importante formule, nous pouvons exprimer les quantités que nous connaissons dans la base canonique en termes de la base locale.

Une autre conséquence du fait que  $e_1, e_2$  et  $e_3$  est une base orthonormée est que, éventuellement en réordonnant les vecteurs, on a

$$\begin{aligned} e_1 \times e_2 &= e_3 \\ e_2 \times e_3 &= e_1 \\ e_3 \times e_1 &= e_2 \end{aligned} \quad (24.79)$$

Ces trois relations s'écrivent en une seule avec

$$e_i \times e_j = \sum_k \epsilon_{ijk} e_k \quad (24.80)$$

où

$$\epsilon_{ijk} = \begin{cases} 0 & \text{si } i, j, k \text{ ne sont pas tous différents} \\ 1 & \text{si } ijk \text{ se ramène à } 123 \text{ par un nombre pair de transpositions} \\ -1 & \text{si } ijk \text{ se ramène à } 123 \text{ par un nombre impair de transpositions} \end{cases} \quad (24.81)$$

est le **symbole de Levi-Civita**. La formule du produit vectoriel peut également être utilisée à l'envers sous la forme

$$e_k = \frac{1}{2} \sum_{ij} \epsilon_{ijk} e_i \times e_j. \quad (24.82)$$

Le symbole de Levi-Civita possède de nombreuses formules. En voici certaines, facilement démontrables en considérant tous les cas :

$$\epsilon_{ijk}\epsilon_{ijl} = \delta_{kl}|\epsilon_{ijk}|. \quad (24.83)$$

Grâce aux symboles de Levi-Civita, le produit mixte des vecteurs de base a une belle forme :

$$e_l \cdot (e_i \times e_j) = \sum_k \epsilon_{ijk} e_l \times e_k = \sum_k \epsilon_{ijk} \delta_{lk} = \epsilon_{ijl}. \quad (24.84)$$

### 24.5.3 Coordonnées polaires

Les coordonnées curvilignes polaires sont données par

$$M(r, \theta) = \begin{pmatrix} r \cos(\theta) \\ r \sin(\theta) \end{pmatrix}, \quad (24.85)$$

et par conséquent

$$\frac{\partial M}{\partial r} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \quad \frac{\partial M}{\partial \theta} = \begin{pmatrix} -r \sin(\theta) \\ r \cos(\theta) \end{pmatrix}. \quad (24.86)$$

Nous avons les normes  $h_r = 1$  et  $h_\theta = r$ , et donc les vecteurs de la base locale en  $(r, \theta)$  sont

$$e_r = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \cos(\theta)e_x + \sin(\theta)e_y \quad (24.87)$$

ainsi que

$$e_\theta = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = -\sin(\theta)e_x + \cos(\theta)e_y. \quad (24.88)$$

Ces vecteurs sont représentés à la figure 24.4. Notez qu'il y en a une paire différente en chaque point.

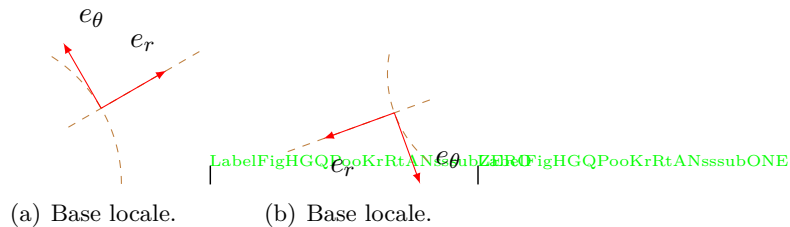


FIGURE 24.4: En brun, les lignes que le point suivrait si on ne variait qu'une coordonnées polaires à la fois. Les vecteurs rouges sont les vecteurs  $e_r$  et  $e_\theta$ .

### 24.5.4 Coordonnées cylindriques

Les coordonnées cylindriques sont les mêmes que les coordonnées polaires à part qu'il faut écrire

$$M(r, \theta, z) = \begin{pmatrix} r \cos(\theta) \\ r \sin(\theta) \\ z \end{pmatrix}, \quad (24.89)$$

et nous avons le vecteur de base supplémentaire

$$e_z = \frac{\partial M}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (24.90)$$

parce que  $h_z = 1$ .

### 24.5.5 Coordonnées sphériques

Les coordonnées curvilignes sphériques sont données par

$$M(\rho, \theta, \varphi) = \begin{pmatrix} \rho \sin(\theta) \cos(\varphi) \\ \rho \sin(\theta) \sin(\varphi) \\ \rho \cos(\theta) \end{pmatrix}, \quad (24.91)$$

dont les dérivées sont données par

$$\begin{aligned} \frac{\partial M}{\partial \rho} &= \begin{pmatrix} \sin(\theta) \cos(\varphi) \\ \sin(\theta) \sin(\varphi) \\ \cos(\theta) \end{pmatrix}, & \frac{\partial M}{\partial \theta} &= \begin{pmatrix} \rho \cos(\theta) \cos(\varphi) \\ \rho \cos(\theta) \sin(\varphi) \\ -\rho \sin(\theta) \end{pmatrix}, \\ \frac{\partial M}{\partial \varphi} &= \begin{pmatrix} -\rho \sin(\theta) \sin(\varphi) \\ \rho \sin(\theta) \cos(\varphi) \\ 0 \end{pmatrix} \end{aligned} \quad (24.92)$$

Les normes de ces vecteurs sont  $h_\rho = 1$ ,  $h_\theta = \rho$  et  $h_\varphi = \rho \sin(\theta)$ . Les vecteurs de la base locale en  $(\rho, \theta, \varphi)$  sont donc

$$\begin{aligned} e_\rho &= \begin{pmatrix} \sin(\theta) \cos(\varphi) \\ \sin(\theta) \sin(\varphi) \\ \cos(\theta) \end{pmatrix}, & e_\theta &= \begin{pmatrix} \cos(\theta) \cos(\varphi) \\ \cos(\theta) \sin(\varphi) \\ -\sin(\theta) \end{pmatrix}, \\ e_\varphi &= \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix} \end{aligned} \quad (24.93)$$

### 24.5.6 Gradient en coordonnées curvilignes

Soit  $(x, y, z) \mapsto f(x, y, z)$  une fonction sur  $\mathbb{R}^3$ . Nous pouvons la composer avec les coordonnées curvilignes  $q$  pour obtenir la fonction

$$\tilde{f}(q_1, q_2, q_3) = f(x_1(q), x_2(q), x_3(q)). \quad (24.94)$$

Nous disons que  $\tilde{f}$  est l'expression de  $f$  dans les coordonnées  $q$ . Nous savons déjà comment calculer le gradient de  $f$  en coordonnées cartésiennes :

$$F(x, y, z) = \nabla f(x, y, z) = \begin{pmatrix} \partial_x f(x, y, z) \\ \partial_y f(x, y, z) \\ \partial_z f(x, y, z) \end{pmatrix}. \quad (24.95)$$

Cela est un vecteur lié au point  $(x, y, z)$ . Notre objectif de bonheur dans la vie serait d'exprimer les coordonnées de ce vecteur dans la base  $\{e_1, e_2, e_3\}$ . En d'autres termes, nous voudrions trouver les nombres  $\tilde{F}_1$ ,  $\tilde{F}_2$  et  $\tilde{F}_3$  tels que

$$F(x, y, z) = F(x(q), y(q), z(q)) = \tilde{F}_1 e_1 + \tilde{F}_2 e_2 + \tilde{F}_3 e_3. \quad (24.96)$$

Ces nombres seront des fonctions de  $(q_1, q_2, q_3)$ .

Par définition,

$$\nabla f = \sum_l \frac{\partial f}{\partial x_l} 1_l. \quad (24.97)$$



En remplaçant  $1_l$  par sa valeur en termes des  $e_i$  par la formule (24.78),

$$\begin{aligned}
 \nabla f &= \sum_l \frac{\partial f}{\partial x_l} 1_l \\
 &= \sum_l \frac{\partial f}{\partial x_l} \sum_i h_i^{-1} \frac{\partial x_l}{\partial q_i} e_i \\
 &= \sum_{il} \frac{1}{h_i} \frac{\partial f}{\partial x_l} \frac{\partial x_l}{\partial q_i} e_i \\
 &= \sum_i \frac{1}{h_i} \frac{\partial \tilde{f}}{\partial q_i} e_i.
 \end{aligned} \tag{24.98}$$

Plus explicitement,

$$\nabla f(x(q), y(q), z(q)) = \sum_i \frac{1}{h_i(q)} \frac{\partial \tilde{f}}{\partial q_i}(q) e_i \tag{24.99}$$

où

$$h_i(q) = \left\| \frac{\partial M}{\partial q_i}(q) \right\|. \tag{24.100}$$

Le plus souvent nous n'allons pas noter explicitement la dépendance de  $h_i$  en  $q$ .

#### 24.5.6.1 Coordonnées sphériques

Nous pouvons exprimer le gradient d'une fonction en coordonnées sphériques en utilisant la formule (24.99) :

$$\nabla \tilde{f}(\rho, \theta, \varphi) = \frac{\partial \tilde{f}}{\partial \rho} e_\rho + \frac{1}{\rho} \frac{\partial \tilde{f}}{\partial \theta} e_\theta + \frac{1}{\rho \sin(\theta)} \frac{\partial \tilde{f}}{\partial \varphi} e_\varphi. \tag{24.101}$$

Cette expression peut paraître peu pratique parce que les vecteurs  $e_\rho$ ,  $e_\theta$  et  $e_\varphi$  eux-mêmes changent en chaque point. Elle est effectivement peu adaptée au dessin, mais elle est très pratique pour des fonctions ayant des symétries.

#### Exemple 24.15.

Le potentiel de la gravitation est la fonction

$$V(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}}. \tag{24.102}$$

En coordonnées sphériques elle s'écrit

$$\tilde{V}(\rho, \theta, \varphi) = \frac{1}{\rho}. \tag{24.103}$$

En voilà une fonction qu'elle est facile à dériver, contrairement à  $V$  ! En suivant la formule (24.101), nous avons immédiatement

$$\nabla \tilde{V} = -\frac{1}{\rho^2} e_\rho. \tag{24.104}$$

Nous voyons immédiatement que cela est un champ de vecteurs dont la norme diminue comme le carré de la distance à l'origine et qui est en permanence dirigé vers l'origine.  $\triangle$

#### 24.5.7 Divergence en coordonnées curvilignes

Nous savons que

$$\nabla \tilde{f} = \sum_j \frac{1}{h_j} \frac{\partial \tilde{f}}{\partial q_j} e_j. \tag{24.105}$$

Nous pouvons en particulier considérer la fonction  $f(q) = q_i$ . De la même manière que nous avons noté  $x_i$  la fonction  $x \mapsto x_i$ , nous notons  $q_i$  la fonction  $q \mapsto q_i$ . Le gradient de cette fonction est donné par

$$\nabla q_i = \sum_j \frac{1}{h_j} \frac{\partial q_i}{\partial q_j} e_j, \quad (24.106)$$

mais  $\frac{\partial q_i}{\partial q_j} = \delta_{ij}$ , donc

$$\nabla q_i = \frac{e_i}{h_i}, \quad (24.107)$$

ou encore

$$e_i = h_i \nabla q_i. \quad (24.108)$$

Cela n'est pas étonnant : la direction dans laquelle la coordonnées  $q_i$  varie le plus est le vecteur  $e_i$  qui donne la tangente à la courbe obtenue lorsque *seul*  $q_i$  varie.

Commençons par calculer la divergence de  $e_i$ . En utilisant la formule (24.82),

$$\nabla \cdot e_k = \frac{1}{2} \sum_{ij} \epsilon_{ijk} \nabla \cdot (e_i \times e_j). \quad (24.109)$$

Nous avons, en utilisant les règles de Leibniz de la proposition 12.547,

$$\begin{aligned} \nabla \cdot (e_i \times e_j) &= \nabla \cdot (h_i \nabla q_i \times h_j \nabla q_j) \\ &= \nabla(h_i h_j) \cdot (\nabla q_i \times \nabla q_j) + h_i h_j \nabla \cdot (\nabla q_i \times \nabla q_j) \\ &= \nabla(h_i h_j) \cdot (\nabla q_i \times \nabla q_j) \\ &\quad + h_i h_j \nabla q_j \cdot \underbrace{(\nabla \times \nabla q_i)}_{=0} \\ &\quad + h_i h_j \nabla q_i \cdot \underbrace{(\nabla \times \nabla q_j)}_{=0} \end{aligned} \quad (24.110)$$

Cela nous fait

$$\nabla \cdot e_k = \sum_{ij} \epsilon_{ijk} \frac{\nabla(h_i h_j)}{h_i h_j} \cdot (e_i \times e_j). \quad (24.111)$$

parce que  $\nabla q_i = h_i^{-1} e_i$ . Nous pouvons développer le gradient qui intervient :

$$\nabla(h_i h_j) = \sum_l \frac{1}{h_l} \frac{\partial}{\partial q_l} (h_i h_j) e_l. \quad (24.112)$$

Nous voyons donc arriver le produit mixte  $e_l \cdot (e_i \times e_j)$ . En utilisant la formule (24.84), cela s'exprime directement sous la forme  $\epsilon_{ijl}$ .

Nous avons alors

$$\begin{aligned} \nabla \cdot e_k &= \frac{1}{2} \sum_{ijl} \frac{1}{h_i h_j h_l} \frac{\partial}{\partial q_l} (h_i h_j) \epsilon_{ijk} \epsilon_{ijl} \\ &= \frac{1}{2} \sum_{ijl} \delta_{kl} |\epsilon_{ijk}| \frac{\partial}{\partial q_l} (h_i h_j) \\ &= \frac{1}{2} \sum_{ij} \frac{|\epsilon_{ijk}|}{h_i h_j h_k} \frac{\partial}{\partial q_k} (h_i h_j). \end{aligned} \quad (24.113)$$

Par exemple,

$$\nabla \cdot e_1 = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial q_1} (h_2 h_3). \quad (24.114)$$

Nous devons maintenant chercher le gradient d'un champ général

$$F(q) = \sum_k F_k(q) e_k. \quad (24.115)$$

La première chose à faire est d'utiliser la formule de Leibniz :

$$\nabla \cdot F = \sum_k \nabla F_k(q) \cdot e_k + \sum_k F_k(q) \nabla \cdot e_k. \quad \text{EqLeibnbablaFek} \quad (24.116)$$

Afin d'alléger les notations, nous allons nous concentrer sur le terme numéro  $k$  et ne pas écrire la somme. Si  $i$  et  $j$  sont les nombres tels que  $\epsilon_{ijk} = 1$ , alors ce que la formule (24.113) signifie, c'est que

$$\nabla \cdot e_k = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial q_k} (h_i h_j). \quad (24.117)$$

Nous savons déjà par la formule (24.99) que

$$\nabla F_k = \sum_l \frac{1}{h_l} \frac{\partial F_k}{\partial q_l} e_l, \quad (24.118)$$

par conséquent

$$\nabla F_k \cdot e_k = \sum_l \frac{1}{h_l} \frac{\partial F_k}{\partial q_l} \delta_{kl} = \frac{1}{h_k} \frac{\partial F_k}{\partial q_k}. \quad (24.119)$$

Pour obtenir cela nous avons utilisé le fait que  $e_l \cdot e_k = \delta_{lk}$ . Le terme numéro  $k$  de la somme (24.116) est donc

$$\frac{1}{h_k} \frac{\partial F_k}{\partial q_k} + \frac{F_k}{h_k h_i h_j} \frac{\partial (h_i h_j)}{\partial q_k} = \frac{1}{h_i h_j h_k} \frac{\partial (F_k h_i h_j)}{\partial q_k} \quad (24.120)$$

où il est entendu que  $i$  et  $j$  représentent les nombres tels que  $\epsilon_{ijk} = 1$ .

Au final, nous avons

$$\nabla \cdot F = \frac{1}{h_1 h_2 h_3} \sum_{ijk} |\epsilon_{ijk}| \frac{\partial (F_k h_i h_j)}{\partial q_k}. \quad (24.121)$$

Ici, la somme sur  $i$  et  $j$  consiste seulement à sélectionner les termes tels que  $i$  et  $j$  ne sont pas  $k$ . En écrivant la somme explicitement,

$$\nabla \cdot F = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial q_1} (F_1 h_2 h_3) + \frac{\partial}{\partial q_2} (F_2 h_1 h_3) + \frac{\partial}{\partial q_3} (F_3 h_1 h_2) \right]. \quad (24.122)$$

#### 24.5.7.1 Coordonnées cylindriques

En coordonnées cylindriques, nous avons déjà vu que  $h_r = 1$ ,  $h_\theta = r$  et  $h_z = 1$ . La divergence est donc donnée par

$$\nabla \cdot F = \frac{1}{r} \left[ \frac{\partial}{\partial r} (r F_r) + \frac{\partial}{\partial \theta} (F_\theta) + \frac{\partial}{\partial z} (r F_z) \right]. \quad \text{EqDivEnCylanf} \quad (24.123)$$

Par exemple si

$$F(r, \theta, z) = r e_\theta + e_z, \quad (24.124)$$

nous avons

$$(\nabla \cdot F)(r, \theta, z) = \frac{1}{r} \left[ \frac{\partial}{\partial \theta} (r) + \frac{\partial}{\partial z} (r) \right] = 0. \quad (24.125)$$

Cela est logique parce que  $r e_\theta$  est à peu près le champ dont nous avons parlé dans l'exemple (12.544), qui était à divergence nulle. En réalité, le champ dont on parlait dans cet exemple était exactement  $-e_\theta$ . Le champ  $e_z$  est également à divergence nulle parce qu'il est constant.

#### 24.5.7.2 Coordonnées sphériques

En coordonnées sphériques, nous avons  $h_\rho = 1$ ,  $h_\theta = r$  et  $h_\varphi = r \sin \theta$ , donc

$$\nabla \cdot F = \frac{1}{r^2 \sin \theta} \left[ \frac{\partial}{\partial \rho} (\rho^2 \sin \theta F_\rho) + \frac{\partial}{\partial \theta} (\rho \sin \theta F_\theta) + \frac{\partial}{\partial \varphi} (\rho F_\varphi) \right]. \quad (24.126)$$

si  $F(\rho, \theta, \varphi) = F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi$ .

### 24.5.8 Laplacien en coordonnées curvilignes orthogonales

Soit une fonction  $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ . Le **laplacien** de  $f$  est donné par

$$\Delta f = \nabla \cdot (\nabla f). \quad (24.127)$$

En utilisant les formules données, nous avons

$$\Delta f = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial q_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial f}{\partial q_1} \right) + \frac{\partial}{\partial q_2} \left( \frac{h_1 h_3}{h_2} \frac{\partial f}{\partial q_2} \right) + \frac{\partial}{\partial q_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial f}{\partial q_3} \right) \right]. \quad (24.128)$$

Dans cette expression, la fonction  $f$  est donnée comme fonction de  $q_1$ ,  $q_2$  et  $q_3$ .

En coordonnées cylindriques, cela s'écrit

$$\begin{aligned} \Delta f &= \frac{1}{r} \left[ \frac{\partial}{\partial r} \left( r \frac{\partial f}{\partial r} \right) + \frac{\partial}{\partial \theta} \left( \frac{1}{r} \frac{\partial f}{\partial \theta} \right) + \frac{\partial}{\partial z} \left( r \frac{\partial f}{\partial z} \right) \right] \\ &= \frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}. \end{aligned} \quad (24.129)$$

Dans cette expression,  $f$  est fonction de  $r$ ,  $\theta$  et  $z$ .

En coordonnées sphériques, cela devient

$$\Delta f = \frac{1}{\rho^2 \sin \theta} \left[ \frac{\partial}{\partial \rho} \left( \rho^2 \sin \theta \frac{\partial f}{\partial \rho} \right) + \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{\partial}{\partial \varphi} \left( \frac{1}{\sin \theta} \frac{\partial f}{\partial \varphi} \right) \right]. \quad (24.130)$$

[EqLaplaceSphe](#)

Dans cette expression,  $f$  est fonction de  $\rho$ ,  $\theta$  et  $\varphi$ .

### 24.5.9 Rotationnel en coordonnées curvilignes orthogonales

Nous voulons calculer le rotationnel de  $F(q) = \sum_k F_k(q) e_k$ . Pour cela nous commençons par écrire  $e_k = h_k \nabla q_k$  et nous utilisons la formule de la proposition [12.547\(3\)](#) avec  $F_k h_k$  en guise de  $f$  :

$$\begin{aligned} \nabla \times F_k e_k &= \nabla \times (F_k h_k \nabla q_k) \\ &= F_k h_k \underbrace{\nabla \times (\nabla q_k)}_{=0} + \nabla (F_k h_k) \times \nabla q_k \\ &= \frac{1}{h_k} \nabla (F_k h_k) \times e_k. \end{aligned} \quad (24.131)$$

Nous utilisons à présent la formule [\(24.99\)](#) du gradient et le formule  $e_j \times e_k = \sum_l \epsilon_{jkl} e_l$  :

$$\begin{aligned} \nabla \times (F_k e_k) &= \sum_j \frac{1}{h_j h_k} \frac{\partial}{\partial q_j} (F_k h_k) e_j \times e_k \\ &= \sum_{jl} \frac{1}{h_j h_k} \epsilon_{jkl} \frac{\partial}{\partial q_j} (F_k h_k) e_l. \end{aligned} \quad (24.132)$$

Le rotationnel s'écrit donc

$$\nabla \times F = \sum_{jkl} \frac{1}{h_j h_k} \epsilon_{jkl} \frac{\partial}{\partial q_j} (F_k h_k) e_l. \quad (24.133)$$

Devant  $e_1$  par exemple nous avons seulement les termes  $j = 2, k = 3$  et  $j = 3, k = 2$ . Étant donné que  $\epsilon_{231} = 1$  et  $\epsilon_{321} = -1$ , le coefficient de  $e_1$  sera simplement

$$\frac{1}{h_2 h_3} \left( \frac{\partial}{\partial q_2} (F_3 h_3) - \frac{\partial}{\partial q_3} (F_2 h_2) \right). \quad (24.134)$$

La formule complète devient

$$\begin{aligned} \nabla \times \sum_k F_k e_k &= \frac{1}{h_2 h_3} \left( \frac{\partial}{\partial q_2} (F_3 h_3) - \frac{\partial}{\partial q_3} (F_2 h_2) \right) e_1 \\ &\quad + \frac{1}{h_1 h_3} \left( \frac{\partial}{\partial q_3} (F_1 h_1) - \frac{\partial}{\partial q_1} (F_3 h_3) \right) e_2 \\ &\quad + \frac{1}{h_2 h_1} \left( \frac{\partial}{\partial q_1} (F_2 h_2) - \frac{\partial}{\partial q_2} (F_1 h_1) \right) e_3. \end{aligned} \quad (24.135)$$

### 24.5.9.1 Coordonnées cylindriques

En utilisant  $h_r = 1$ ,  $h_\theta = r$  et  $h_z = 1$ , nous trouvons

$$\begin{aligned}\nabla \times (F_r e_r + F_\theta e_\theta + F_z e_z) &= \frac{1}{r} \left( \frac{\partial F_z}{\partial \theta} - \frac{\partial (F_\theta r)}{\partial z} \right) e_r \\ &+ \left( \frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r} \right) e_\theta \\ &+ \left( \frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta} \right) e_z.\end{aligned}\tag{24.136}$$

EqRotationnelCylind

### 24.5.9.2 Coordonnées sphériques

En utilisant  $h_\rho = 1$ ,  $h_\theta = \rho$  et  $h_\varphi = \rho \sin \theta$ , nous trouvons

$$\begin{aligned}\nabla \times (F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi) &= \frac{1}{\rho \sin \theta} \left( \frac{\partial (F_\varphi) \sin \theta}{\partial \theta} - \frac{\partial F_\theta}{\partial \varphi} \right) e_\rho \\ &+ \frac{1}{\rho \sin \theta} \left( \frac{\partial F_\rho}{\partial \varphi} - \frac{\partial (F_\varphi \rho \sin \theta)}{\partial \rho} \right) e_\theta \\ &+ \frac{1}{\rho} \left( \frac{\partial F_\theta \rho}{\partial \rho} - \frac{\partial F_r}{\partial \theta} \right) e_\varphi.\end{aligned}\tag{24.137}$$

Note : dans le premier terme, il y a une simplification par  $\rho$ .

## 24.6 Les formules

### 24.6.1 Coordonnées polaires

Les vecteurs de base :

$$e_r = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \cos(\theta) e_x + \sin(\theta) e_y \tag{24.138a}$$

$$e_\theta = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = -\sin(\theta) e_x + \cos(\theta) e_y. \tag{24.138b}$$

Le gradient :

$$\nabla \tilde{f}(r, \theta) = \frac{\partial \tilde{f}}{\partial r}(r, \theta) e_r + \frac{1}{r} \frac{\partial \tilde{f}}{\partial \theta}(r, \theta) e_\theta. \tag{24.139}$$

La divergence :

$$\nabla \cdot F = \frac{1}{r} \left[ \frac{\partial}{\partial r}(r F_r) + \frac{\partial}{\partial \theta}(F_\theta) \right]. \tag{24.140}$$

EqBxJKd

Le rotationnel :

$$\nabla \times (F_r e_r + F_\theta e_\theta) = \left( \frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta} \right) e_z. \tag{24.141}$$

EqBpqCw

Notons que le rotationnel n'existe pas vraiment en deux dimensions. Ici nous avons vu le champ  $F(r, \theta)$  comme un champ dans  $\mathbb{R}^3$  ne dépendant pas de  $z$  et n'ayant pas de composante  $z$ . Le résultat est un rotationnel qui est dirigé selon l'axe  $z$ .

### 24.6.2 Coordonnées cylindriques

Les vecteurs de base : idem qu'en coordonnées polaires, et on ajoute  $e_z$  sans modifications.

Le gradient :

$$\nabla \tilde{f}(r, \theta, z) = \frac{\partial \tilde{f}}{\partial r}(r, \theta, z) e_r + \frac{1}{r} \frac{\partial \tilde{f}}{\partial \theta}(r, \theta, z) e_\theta + \frac{\partial \tilde{f}}{\partial z}(r, \theta, z) e_z. \tag{24.142}$$

La divergence :

$$\nabla \cdot F = \frac{1}{r} \left[ \frac{\partial}{\partial r}(rF_r) + \frac{\partial}{\partial \theta}(F_\theta) + \frac{\partial}{\partial z}(rF_z) \right]. \quad (24.143)$$

Le rotationnel :

$$\begin{aligned} \nabla \times (F_r e_r + F_\theta e_\theta + F_z e_z) &= \frac{1}{r} \left( \frac{\partial F_z}{\partial \theta} - \frac{\partial(F_\theta r)}{\partial z} \right) e_r \\ &+ \left( \frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r} \right) e_\theta \\ &+ \left( \frac{\partial(F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta} \right) e_z. \end{aligned} \quad (24.144)$$

Note : les formules concernant les coordonnées polaires se réduisent de celles-ci en enlevant toutes les références à  $z$ .

### 24.6.3 Coordonnées sphériques

Les vecteurs de base :

$$\begin{aligned} e_r &= \begin{pmatrix} \sin(\theta) \cos(\varphi) \\ \sin(\theta) \sin(\varphi) \\ \cos(\theta) \end{pmatrix}, \quad e_\theta = \begin{pmatrix} \cos(\theta) \cos(\varphi) \\ \cos(\theta) \sin(\varphi) \\ -\sin(\theta) \end{pmatrix}, \\ e_\varphi &= \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \\ 0 \end{pmatrix} \end{aligned} \quad (24.145)$$

Le gradient :

$$\nabla \tilde{f}(\rho, \theta, \varphi) = \frac{\partial \tilde{f}}{\partial \rho} e_\rho + \frac{1}{\rho} \frac{\partial \tilde{f}}{\partial \theta} e_\theta + \frac{1}{\rho \sin(\theta)} \frac{\partial \tilde{f}}{\partial \varphi} e_\varphi. \quad (24.146)$$

La divergence :

$$\nabla \cdot F = \frac{1}{\rho^2 \sin \theta} \left[ \frac{\partial}{\partial \rho}(\rho^2 \sin \theta F_\rho) + \frac{\partial}{\partial \theta}(\rho \sin \theta F_\theta) + \frac{\partial}{\partial \varphi}(\rho F_\varphi) \right]. \quad (24.147)$$

Le rotationnel :

$$\begin{aligned} \nabla \times (F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi) &= \frac{1}{\rho \sin \theta} \left( \frac{\partial(F_\varphi) \sin \theta}{\partial \theta} - \frac{\partial F_\theta}{\partial \varphi} \right) e_\rho \\ &+ \frac{1}{\rho \sin \theta} \left( \frac{\partial F_\rho}{\partial \varphi} - \frac{\partial(F_\varphi \rho \sin \theta)}{\partial \rho} \right) e_\theta \\ &+ \frac{1}{\rho} \left( \frac{\partial F_\theta \rho}{\partial \rho} - \frac{\partial F_\rho}{\partial \theta} \right) e_\varphi. \end{aligned} \quad (24.148)$$



# Chapitre 25

## Espaces de Hilbert

### 25.1 Espaces de Hilbert

#### Définition 25.1.

Soit un espace vectoriel  $E$  sur le corps  $\mathbb{K}$ . Son **dual topologique**, noté  $E'$ , est l'ensemble des formes linéaires continues de  $E$  vers  $\mathbb{K}$ .

#### Proposition 25.2 ([684]).

Soient des espaces vectoriels normés  $X$  et  $Y$  ainsi qu'une forme sesquilinéaire  $\phi: X \times Y \rightarrow \mathbb{C}$ . L'espace  $X \times Y$  est muni de la topologie produit<sup>1</sup>. Il y a équivalence des faits suivants.

- (1)  $\phi$  est continue.
- (2)  $\phi$  est continue en  $(0, 0)$
- (3)  $\sup\{|\phi(x, y)| \mid \text{tel que } \|x\|, \|y\| \leq 1\} < \infty$ .
- (4) Il existe  $C \geq 0$  telle que

$$|\phi(x, y)| \leq C\|x\|\|y\| \quad (25.1)$$

pour tout  $(x, y) \in X \times Y$ .

- (5)  $\phi$  est bornée

Dans le cas où  $\phi$  vérifie ces propriétés, alors nous posons

$$\mathcal{A} = \{|\phi(x, y)| \mid \text{tel que } \|(x, y)\| \leq 1\} \quad (25.2)$$

et

$$\mathcal{B} = \{\beta \geq 0 \mid \text{tel que } |\phi(x, y)| \leq \beta\|x\|\|y\|, \forall x, y\} \quad (25.3)$$

et nous avons

$$\|\phi\| = \sup(\mathcal{A}) = \min(\mathcal{B}). \quad (25.4)$$

*Démonstration.* En plusieurs parties.

- (1) **(1)  $\Rightarrow$  (2)** Une application est continue si et seulement si elle est continue en tout point (théorème 7.262).
- (2) **(2)  $\Rightarrow$  (3)** Pour rappel, nous avons mis sur  $X \times Y$  la topologie produit. Le lemme 7.282 nous indique que cette topologie est la même que la topologie de la norme produit. Bref, nous supposons que  $\phi$  est continue en  $(0, 0)$  et nous prouvons que

$$\sup\{|\phi(x, y)| \mid \text{tel que } \|x\|, \|y\| \leq 1\} < \infty. \quad (25.5)$$

Nous y allons par contradiction. Supposons que pour  $n$  dans  $\mathbb{N}$  il existe  $(x_n, y_n) \in X \times Y$  tels que  $\|x_n\|, \|y_n\| \leq 1$  et  $|\phi(x_n, y_n)| > n^2$ . En posant  $v_n = x_n/n$  et  $w_n = y_n/n$ , nous avons

$$\|v_n\| \leq \frac{\|x_n\|}{n} \leq \frac{1}{n}, \quad (25.6)$$

---

1. Quand on ne précise pas, un produit est toujours muni de la topologie produit 7.63.



et la même chose pour  $\|w_n\|$ . Nous avons  $(v_n, w_n) \xrightarrow{X \times Y} (0, 0)$  et donc, par continuité de  $\phi$  en  $(0, 0)$  nous avons  $\phi(v_n, w_n) \xrightarrow{\mathbb{C}} \phi(0, 0) = 0$ . Mais

$$|\phi(v_n, w_n)| = \frac{1}{n^2} |\phi(x_n, y_n)| \geq \frac{1}{n^2} n^2 = 1. \quad (25.7)$$

Donc nous n'avons pas  $|\phi(v_n, w_n)| \rightarrow 0$ .

(3) (3)  $\Rightarrow$  (4) Nous notons

$$M = \sup\{|\phi(x, y)| \text{ tel que } \|x\|, \|y\| \leq 1\} < \infty, \quad (25.8)$$

et nous allons prouver que  $|\phi(x, y)| \leq M\|x\|\|y\|$ .

Si  $x = 0$  ou si  $y = 0$ , c'est évident ; nous supposons donc que  $x$  et  $y$  sont tous deux non nuls. Nous posons  $v = x/\|x\|$  et  $w = y/\|y\|$ . Nous avons alors

$$|\phi(x, y)| = \|x\|\|y\||\phi(v, w)| \leq \|x\|\|y\|M. \quad (25.9)$$

(4) (4)  $\Rightarrow$  (5) La norme de  $\phi$  est le nombre

$$\|\phi\| = \sup_{\substack{(x,y) \in X \times Y \\ \|(x,y)\|=1}} |\phi(x, y)|, \quad (25.10)$$

que nous devons prouver être fini. Les éléments  $(x, y)$  sur lesquels porte le supremum vérifient

$$1 = \|(x, y)\| = \max\{\|x\|, \|y\|\}, \quad (25.11)$$

et donc vérifient  $\|x\|, \|y\| \leq 1$ . En utilisant (4), nous avons donc

$$\|\phi\| = \sup_{\|(x,y)\|=1} |\phi(x, y)| \leq \sup_{\|(x,y)\|=1} C\|x\|\|y\| \leq C. \quad (25.12)$$

(5) (5)  $\Rightarrow$  (1) Il nous reste à prouver que si  $\phi$  est bornée, alors elle est continue. Soit une suite convergente  $(x_n, y_n) \xrightarrow{X \times Y} (x, y)$ . Nous allons prouver que  $\phi(x_n, y_n) \xrightarrow{\mathbb{C}} \phi(x, y)$ . Ce sont les majorations usuelles :

$$|\phi(x_n, y_n) - \phi(x, y)| \leq |\phi(x_n, y_n) - \phi(x_n, y)| + |\phi(x_n, y) - \phi(x, y)| \quad (25.13a)$$

$$= |\phi(x_n, y_n - y)| + |\phi(x_n - x, y)| \quad (25.13b)$$

$$\leq \|\phi\|\|x_n\|\|y_n - y\| + \|\phi\|\|x_n - x\|\|y\| \quad (25.13c)$$

$$\rightarrow 0. \quad (25.13d)$$

Enfin nous supposons que toutes ces conditions équivalentes, et nous prouvons que  $\|\phi\| = \sup(\mathcal{A}) = \min(\mathcal{B})$ .

(1)  $\|\phi\| = \sup(\mathcal{A})$  En posant

$$\mathcal{C} = \{|\phi(x, y)| \text{ tel que } \|(x, y)\| = 1\}, \quad (25.14)$$

nous avons, par définition,  $\|\phi\| = \sup(\mathcal{C})$ . Étant donné que  $\mathcal{C} \subset \mathcal{A}$ , nous avons  $\sup(\mathcal{C}) \leq \sup(\mathcal{A})$ . Faisons l'inégalité dans l'autre sens. Soit  $|\phi(x, y)| \in \mathcal{A}$ , c'est-à-dire  $\|(x, y)\| \leq 1$ . En posant  $v = x/\|x\|$  et  $w = y/\|y\|$  nous avons  $|\phi(v, w)| \in \mathcal{C}$  et surtout  $\|x\|, \|y\| \leq 1$ . Donc

$$|\phi(v, w)| = \frac{1}{\|x\|\|y\|} |\phi(x, y)| \geq |\phi(x, y)|. \quad (25.15)$$

Donc tout élément de  $\mathcal{C}$  est majoré par un élément de  $\mathcal{A}$ . Donc  $\sup(\mathcal{A}) \geq \sup(\mathcal{C})$ .

(2)  $\|\phi\| \in \mathcal{B}$  Soit  $(x, y) \in X \times Y$ . Étant donné que  $\|x/\|x\|\| = 1$  nous avons

$$\left| \phi \left( \frac{x}{\|x\|}, \frac{y}{\|y\|} \right) \right| \leq \sup_{\|(v,w)\|=1} |\phi(v, w)| = \|\phi\|. \quad (25.16)$$

En multipliant les deux côtés par  $\|x\|\|y\|$  nous trouvons

$$|\phi(x, y)| \leq \|\phi\| \|x\| \|y\|, \quad (25.17)$$

et donc  $\|\phi\| \in \mathcal{B}$ .

(3)  $\|\phi\|$  **minore**  $\mathcal{B}$

Nous montrons que pour tout  $\alpha \in \mathcal{A}$  et pour tout  $\beta \in \mathcal{B}$ , nous avons  $\alpha \leq \beta$ . En effet si  $\alpha \in \mathcal{A}$ , alors  $\alpha = |\phi(x, y)|$  pour un certain  $(x, y) \in X \times Y$  vérifiant  $\|(x, y)\| = 1$ . Mais  $\beta \in \mathcal{B}$  signifie que  $|\phi(x, y)| \leq \beta \|x\| \|y\|$ . Nous avons donc

$$\alpha = |\phi(x, y)| \leq \beta \|x\| \|y\| \leq \beta \quad (25.18)$$

parce que  $\|x\|, \|y\| \leq 1$ .

□

DefORuBdBN

### Définition 25.3.

Un espace **préhilbertien** est

- soit un espace vectoriel réel muni d'un produit scalaire<sup>2</sup>,
- soit un espace vectoriel complexe muni d'un produit scalaire hermitien<sup>3</sup>.

Un **espace de Hilbert** est un espace préhilbertien qui est complet<sup>4</sup> pour la norme induite par son produit (scalaire ou hermitien), c'est-à-dire tel que

$$\|x\| = \sqrt{\langle x, x \rangle} \quad (25.19)$$

pour tout élément  $x$ .

Dans les deux cas nous considérons la topologie métrique dérivant du produit.

Dans les cas de dimension finie, les espaces vectoriels normés sont automatiquement complets par la proposition 7.345.

La différence entre un espace de Hilbert et un espace de Banach est que dans le cas d'un espace de Hilbert, nous demandons que la norme dérive d'un produit scalaire.

PropTdupIG

### Proposition 25.4.

Si  $\mathcal{H}$  est un espace de Hilbert réel, alors

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle. \quad (25.20)$$

Si  $\mathcal{H}$  est un espace de Hilbert complexe alors

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2 \operatorname{Re} \langle x, y \rangle. \quad (25.21)$$

EgrbBlkK

Dans les deux cas nous avons l'inégalité de Cauchy-Schwarz :

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \quad (25.22)$$

Dans un espace vectoriel de dimension infinie, tous les opérateurs linéaires ne sont pas continus.

2. Définition 8.164.

3. Définition 8.172.

4. Définition 7.310.

**Exemple 25.5.**

Soit un espace vectoriel  $V$  engendré par la base  $\{e_k\}_{k \in \mathbb{N}}$  et l'application linéaire  $T: V \rightarrow V$  donnée par

$$Te_k = ke_k. \quad (25.23)$$

Nous allons montrer que l'image inverse de la boule unité ouverte  $\mathcal{O}$  n'est pas ouverte. En effet si  $(\epsilon_k)$  est une suite de réels strictement positifs tendant vers zéro, les vecteurs

$$a_k = \left(\frac{1}{k} + \epsilon_k\right) e_k \quad (25.24)$$

sont hors de  $T^{-1}\mathcal{O}$  parce que

$$Ta_k = (1 + k\epsilon_k)e_k. \quad (25.25)$$

Mais la suite  $(a_k)$  converge vers 0 qui fait partie de  $T^{-1}\mathcal{O}$ . Donc le complémentaire de  $T^{-1}\mathcal{O}$  n'est pas fermé, ce qui prouve que  $T^{-1}\mathcal{O}$  n'est pas ouvert.  $\triangle$

PROPOoMSAYooONHLYq

**Proposition 25.6 ([1]).**

Soit un espace de Hilbert  $\mathcal{H}$  et un espace vectoriel normé  $(E, \|\cdot\|)$  tel qu'il existe une bijection linéaire isométrique  $\Phi: \mathcal{H} \rightarrow E$ . Alors en posant

$$\langle \alpha, \beta \rangle_E = \langle \Phi^{-1}(\alpha), \Phi^{-1}(\beta) \rangle_{\mathcal{H}}, \quad (25.26)$$

l'espace  $E$  devient un espace de Hilbert.

*Démonstration.* Le fait que  $\langle \cdot, \cdot \rangle_E$  soit un produit scalaire (hermitien) lorsque  $\langle \cdot, \cdot \rangle_{\mathcal{H}}$  l'est est une conséquence de la linéarité de  $\Phi$ . Il donne la norme parce que

$$\sqrt{\langle \alpha, \alpha \rangle} = \sqrt{\langle \Phi^{-1}(\alpha), \Phi^{-1}(\alpha) \rangle} = \|\Phi^{-1}(\alpha)\| = \|\alpha\| \quad (25.27)$$

parce que  $\Phi^{-1}$  est une isométrie.

Pour voir que  $E$  est complet, soit une suite  $(\alpha_k)$  de  $\tau$ -Cauchy dans  $E$ . Nous devons prouver qu'elle converge. Nous posons  $x_k = \Phi^{-1}(\alpha_k)$ .

Commençons par voir que  $(x_k)$  est  $\tau$ -Cauchy dans  $\mathcal{H}$ . Soit un voisinage  $V$  de 0 dans  $\mathcal{H}$ . Nous posons  $U = \Phi(V)$ . Vu que  $(\alpha_k)$  est  $\tau$ -Cauchy, il existe  $N > 0$  tel que  $\alpha_k - \alpha_l \in U$  dès que  $k, l > N$ . Du coup, si  $k, l > N$  nous avons aussi  $x_k - x_l = \Phi^{-1}(\alpha_k - \alpha_l) \in \Phi^{-1}(U) = V$ . Cela prouve que  $(x_k)$  est  $\tau$ -Cauchy.

Vu que  $\mathcal{H}$  est complet, il existe  $x \in \mathcal{H}$  tel que  $x_k \xrightarrow{\mathcal{H}} x$ . Montrons que  $\alpha_k \xrightarrow{E} \Phi(x)$ .

Soit  $\epsilon > 0$  et  $N > 0$  tel que  $\|x_k - x\| < \epsilon$  dès que  $k \geq N$ . Alors nous avons aussi

$$\|\alpha_k - \Phi(x)\| = \|\Phi(x_k) - \Phi(x)\| = \|\Phi(x_k - x)\| = \|x_k - x\| < \epsilon \quad (25.28)$$

où nous avons utilisé le fait que  $\Phi$  était linéaire et une isométrie.  $\square$

**25.1.1 Sous-espace vectoriel fermé ???**

Nous verrons que beaucoup de résultats demandent un sous-espace vectoriel *fermé*. Une question légitime est : est-ce qu'il existe des sous-espaces vectoriels qui ne soient pas fermés ? En dimension finie, tous les sous-espaces vectoriels sont fermés<sup>5</sup>, mais cela n'est pas vrai en dimension infinie.

Soit en effet une partie libre infinie  $A = \{v_i\}_{i \in \mathbb{N}}$  dans un espace de Hilbert  $\mathcal{H}$ . L'ensemble  $\text{Span}(A)$  des combinaisons linéaires d'éléments de  $A$  est un sous-espace vectoriel de  $\mathcal{H}$ , mais il n'est pas fermé.

En effet, supposons pour simplifier les notations que  $\|v_i\| = 1$  pour tout  $i$ . Alors nous considérons la combinaison

$$a = \sum_{k=1}^{\infty} \alpha_k v_k \quad (25.29)$$

---

5. Proposition 7.229.

où les  $\alpha_k$  sont suffisamment décroissants pour assurer les convergences<sup>6</sup>. Le vecteur  $a$  n'est pas dans  $\text{Span}(A)$ , mais la suite  $a_n = \sum_{k=1}^n \alpha_k v_k$  est dans  $\text{Span}(A)$  et converge vers  $a$  :  $a_n \xrightarrow{\mathcal{H}} a$ . En effet,

$$\|a_n - a\| = \left\| \sum_{k=n+1}^{\infty} \alpha_k v_k \right\| \leq \sum_{k=n+1}^{\infty} |\alpha_k| \xrightarrow{n \rightarrow \infty} 0. \quad (25.30)$$

Vous voulez des détails sur la dernière limite ? Vu que la somme  $\sum_k |\alpha_k|$  converge, la suite des sommes de queues de suites<sup>7</sup> converge vers zéro :

$$\lim_{n \rightarrow \infty} \sum_{k=n+1}^{\infty} |\alpha_k| = 0. \quad (25.31)$$

## 25.2 Théorème de la projection

Voir la version plus simple dans 18.104.

ThoProjOrthuZeYkz

**Théorème 25.7** (Projection sur partie fermée convexe[685, 686]).

Soit  $\mathcal{H}$  un espace de Hilbert,  $x \in \mathcal{H}$ , et  $C$  un sous-ensemble fermé convexe de  $\mathcal{H}$ .

(1) Les deux conditions suivantes sur  $y \in \mathcal{H}$  sont équivalentes :

ETsfYCSItemi

(1a)  $\|x - y\| = \inf\{\|x - z\| \text{ tel que } z \in C\}$ ,

ETsfYCSItemii

(1b) pour tout  $z \in C$ ,  $\text{Re}\langle x - y, z - y \rangle \leq 0$ .

(2) Il existe un unique  $y \in \mathcal{H}$ , noté  $y = \text{proj}_C(x)$  vérifiant ces conditions.

*Démonstration.* Nous commençons par prouver l'existence et l'unicité d'un élément dans  $C$  vérifiant la première condition. Ensuite nous verrons l'équivalence.

Nous nommons  $d$  l'infimum en question de la première condition.

**Existence** Soit  $(y_n)$  une suite dans  $C$  telle que

$$\lim_{n \rightarrow \infty} \|x - y_n\| = \inf\{\|x - z\| \text{ tel que } z \in C\} = d. \quad (25.32)$$

Nous allons montrer que cette suite peut être choisie de Cauchy. Elle convergera donc dans  $\mathcal{H}$  parce que ce dernier est complet. Mais  $C$  étant supposé fermé dans  $\mathcal{H}$ , la limite appartiendra à  $C$ . Soient  $r, s \in \mathbb{N}$ . D'abord nous avons

$$\|y_r - y_s\|^2 = \langle y_r - y_s + x - x, y_r - y_s + x - x \rangle \quad (25.33a)$$

$$= \|y_r - x\|^2 + \|y_s - x\|^2 - 2\langle y_r - x, y_s - x \rangle. \quad (25.33b)$$

Ensuite,

$$4 \left\| \frac{y_r + y_s}{2} - x \right\|^2 = \langle y_r + y_s - 2x, y_r + y_s - 2x \rangle \quad (25.34a)$$

$$= \|y_r - x\|^2 + \|y_s - x\|^2 + 2\langle y_r - x, y_s - x \rangle. \quad (25.34b)$$

Si nous égalisons les valeurs de  $2\langle y_r - x, y_s - x \rangle$  nous trouvons

$$\|y_r - y_s\|^2 = -4 \left\| \frac{y_r + y_s}{2} - x \right\|^2 + 2\|y_r - x\|^2 + 2\|y_s - x\|^2. \quad (25.35)$$

EoigCyUa

La distance infimum étant  $d$ , nous pouvons choisir  $y_n$  de telle façon à avoir

$$\|y_n - x\| \leq d + \frac{1}{n}. \quad (25.36)$$

6. Par exemple  $\alpha_k = 1/k^2$  si les  $v_k$  sont orthonormaux.

7. On se comprend hein.

D'autre part étant donné que  $C$  est convexe,  $(y_r + y_s)/2$  est dans  $C$  et nous avons

$$\left\| \frac{y_r + y_s}{2} - x \right\| \geq d. \quad (25.37)$$

En mettant ces majorations dans (25.35) nous trouvons

$$\|y_r - y_s\|^2 \leq -4d + 2 \left( d + \frac{1}{r} \right) + 2 \left( d + \frac{1}{s} \right) = \frac{1}{r} + \frac{1}{s}. \quad (25.38)$$

La suite  $(y_n)$  est donc de Cauchy et la limite est un élément de  $C$ . Prouvons que cet élément  $y$  réalise l'infimum. Pour cela nous avons les inégalités

$$d \leq \|x - y\| \leq \|x - y_n\| + \|y_n - y\|. \quad (25.39)$$

En prenant le limite  $n \rightarrow \infty$  nous trouvons

$$d \leq \|x - y\| \leq d. \quad (25.40)$$

**Unicité** Même preuve que pour le théorème en dimension finie 12.150.

(1a)  $\Rightarrow$  (1b) Même preuve que pour le théorème en dimension finie 12.150.

(1b)  $\Rightarrow$  (1a) Même preuve que pour le théorème en dimension finie 12.150.

□

PropAXJpCe

### Proposition 25.8.

Soit  $C$  une partie convexe et fermée de l'espace de Hilbert  $\mathcal{H}$ . Alors pour tout  $x_1, x_2 \in \mathcal{H}$  nous avons

$$\| \text{proj}_C(x_1) - \text{proj}_C(x_2) \| \leq \|x_1 - x_2\|. \quad (25.41)$$

En particulier la projection est une application continue.

*Démonstration.* Nous posons  $y_1 = \text{proj}_C(x_1)$  et  $y_2 = \text{proj}_C(x_2)$ . Par la partie (1b) du théorème 25.7, nous avons, pour tout  $z, z' \in C$  les inégalités

$$\text{Re} \langle x_1 - y_1, z - y_1 \rangle \leq 0 \quad (25.42a)$$

$$\text{Re} \langle x_2 - y_2, z' - y_2 \rangle \leq 0. \quad (25.42b)$$

En prenant  $z = y_2$  et  $z' = y_1$  et en sommant nous trouvons

$$\text{Re} \langle (x_1 - y_1) + (y_2 - x_2), y_2 - y_1 \rangle \leq 0. \quad (25.43)$$

Nous pouvons maintenant calculer

$$\begin{aligned} \|y_1 - y_2\|^2 &= \text{Re} \|y_1 - y_2\|^2 \\ &= \text{Re} \langle y_1 - y_2, (y_1 - x_1) + x_1 - x_2 + (x_2 - y_2) \rangle \\ &= \text{Re} \langle y_1 - y_2, x_1 - x_2 \rangle + \text{Re} \langle y_2 - y_1, (x_1 - y_1) + (y_2 - x_2) \rangle \\ &\leq \langle y_1 - y_2, x_1 - x_2 \rangle \\ &\leq \| \langle y_1 - y_2, x_1 - x_2 \rangle \| \\ &\leq \|y_1 - y_2\| \|x_1 - x_2\|. \end{aligned} \quad (25.44)$$

En simplifiant par  $\|y_1 - y_2\|$ <sup>8</sup> nous trouvons le résultat

$$\|y_1 - y_2\| \leq \|x_1 - x_2\|. \quad (25.45)$$

□

---

8. Si c'est nul, alors la preuve est évidente.

ThoMXwOjb

**Théorème 25.9** (Projection orthogonale).

Soit  $\mathcal{H}$  un espace de Hilbert et  $K$ , un sous-espace vectoriel fermé non réduit à  $\{0\}$  et  $x \in \mathcal{H}$ . L'élément  $y = \text{proj}_K(x)$  est l'unique élément de  $K$  tel que

$$x - y \in K^\perp. \quad (25.46)$$

De plus l'application  $x \mapsto \text{proj}_K(x)$  est linéaire, continue et de norme 1.

L'élément  $y$  ainsi défini est la **projection orthogonale** de  $x$  sur  $K$  et sera noté  $\text{proj}_K(f)$ .

*Démonstration.* La continuité de la projection est donnée par la proposition 25.8.

Soit  $z$  un élément de  $K$  tel que  $\langle z - x, a \rangle = 0$  pour tout  $a \in K$ . Nous avons

$$\|x - a\|^2 = \|z - x\|^2 + \|a - z\|^2 + 2 \underbrace{\langle z - x, a - z \rangle}_{=0} \quad (25.47a)$$

$$\geq \|z - x\|^2. \quad (25.47b)$$

Le produit scalaire est nul parce que  $a - z \in K$ . La distance  $\|z - x\|$  est donc bien la plus petite distance entre  $x$  et les éléments de  $K$ .

Dans l'autre sens, nous supposons que  $y \in K$  minimise la distance à  $x$  dans  $K$ . Par hypothèse pour tout  $a$  et pour tout  $\lambda \in \mathbb{R}$ , la différence

$$\|(y + \lambda a) - x\|^2 - \|y - x\|^2 \quad (25.48)$$

est positive. En développant les produits scalaires nous trouvons la conditions suivante

$$\lambda^2 \|a\|^2 + 2\lambda \langle a, y - x \rangle \geq 0 \quad (25.49)$$

qui doit être vraie pour tout  $\lambda \in \mathbb{R}$ . En tant que polynôme du second degré en  $\lambda$ , cela n'aura pas deux racines réelles distinctes uniquement si  $\langle a, y - x \rangle = 0$ .

Nous montrons maintenant la linéarité de la projection orthogonale. Soient  $x_1, x_2 \in \mathcal{H}$ . L'élément  $y = \text{proj}_K x_1 + \text{proj}_K x_2$  satisfait à la condition d'orthogonalité : pour tout  $z \in K$ ,

$$\langle x_1 + x_2 - \text{proj}_K x_1 - \text{proj}_K x_2, z \rangle = \langle x_1 - \text{proj}_K x_1, z \rangle + \langle x_2 - \text{proj}_K x_2, z \rangle = 0. \quad (25.50)$$

Étant donné que  $K$  est un sous-espace vectoriel, la condition de minimalité est automatiquement vérifiée (seconde partie du théorème 25.7).

En ce qui concerne la norme opérateur de  $\text{proj}_K$ , la décomposition de  $x \in \mathcal{H}$  en composantes dans  $K$  et  $K^\perp$  est

$$x = \text{proj}_K x + (x - \text{proj}_K x). \quad (25.51)$$

Étant deux parties orthogonales nous avons

$$\|\text{proj}_K x\|^2 = \|x\|^2 - \|x - \text{proj}_K x\|^2. \quad (25.52)$$

En prenant  $\|x\| = 1$  nous trouvons  $\|\text{proj}_K x\|^2 \leq 1$  et par conséquent  $\|\text{proj}_K\| \leq 1$ . Mais d'autre part en prenant  $x \in K$  nous avons automatiquement  $\|\text{proj}_K\| \geq 1$ .  $\square$

**Proposition 25.10.**

Soit  $\mathcal{H} = L^2(\Omega, \mathcal{A}, \mu)$ ,  $\mathcal{F}$  une sous tribu de  $\mathcal{A}$  et  $K$  l'ensemble de fonctions  $\mathcal{F}$ -mesurables dans  $L^2(\Omega, \mathcal{A}, \mu)$ . Si  $f \in L^2(\Omega, \mathcal{A}, \mu)$  est positive, alors  $\text{proj}_K f$  est positive (presque partout).

*Démonstration.* L'ensemble  $A = \{\text{proj}_K f < 0\}$  est dans  $\mathcal{F}$ . En effet

$$A = (\text{proj}_K f)^{-1}([-\infty, 0]) \quad (25.53)$$

alors que, par construction,  $\text{proj}_K f$  est  $\mathcal{F}$ -mesurable. La fonction indicatrice  $\mathbb{1}_A$  est alors  $\mathcal{F}$ -mesurable (c'est-à-dire  $\mathbb{1}_A \in K$ ) et nous avons

$$0 \leq \int_\Omega f \mathbb{1}_A = \int_\Omega \text{proj}_K f \mathbb{1}_A \leq 0. \quad (25.54)$$

Étant donné que nous avons supposé  $f \geq 0$  nous avons alors  $\mu(A) = 0$ . D'où le fait que  $\text{proj}_K f$  est presque partout positive.  $\square$

### 25.3 Systèmes orthogonaux et bases

Dans cette partie nous noterons  $\mathbb{K}$  le corps de base de l'espace  $\mathcal{H}$ . Seuls deux cas sont envisagés :  $\mathbb{K} = \mathbb{C}$  ou  $\mathbb{K} = \mathbb{R}$ .

Pour chaque  $x \in \mathcal{H}$  nous considérons l'application

$$\begin{aligned} \Phi_x: \mathcal{H} &\rightarrow \mathbb{K} \\ x &\mapsto \langle x, x \rangle. \end{aligned} \tag{25.55}$$

Ce sont des applications continues.

#### 25.3.1 Orthogonal d'une partie

##### Définition 25.11.

Soit une partie  $A$  de l'espace de Hilbert  $\mathcal{H}$ . L'**orthogonal** de  $A$  est l'ensemble

$$A^\perp = \{v \in \mathcal{H} \text{ tel que } \langle v, x \rangle = 0 \forall x \in A\}. \tag{25.56}$$

##### Proposition 25.12.

Si  $\mathcal{H}$  est un préhilbert et si  $A \subset \mathcal{H}$ , alors l'ensemble  $A^\perp$  est un sous-espace fermé de  $\mathcal{H}$ .

*Démonstration.* L'application  $\Phi_x$  définie en (25.55) est continue pour chaque  $x \in \mathcal{H}$  et par conséquent l'ensemble  $\ker \Phi_x = \Phi_x^{-1}(\{0\})$  est fermé. L'ensemble

$$A^\perp = \bigcap_{x \in A} \ker \Phi_x \tag{25.57}$$

est donc fermé. □

##### Théorème 25.13.

Soit  $\mathcal{H}$  un espace de Hilbert et  $F$  un sous-espace fermé de  $\mathcal{H}$ . Alors

(1) Nous avons la décomposition

$$\mathcal{H} = F \oplus F^\perp. \tag{25.58}$$

(2) La projection sur  $F$  par rapport à la somme directe (25.58) est la projection  $\text{proj}_F$  du théorème de projection.

(3) La décomposition (25.58) est topologique.

##### Proposition 25.14.

Si  $F$  est un sous-espace de l'espace de Hilbert  $\mathcal{H}$  alors on a  $F^{\perp\perp} = \bar{F}$ .

*Démonstration.* Nous savons par la proposition 25.12 que  $F^\perp$  est fermé, par conséquent le théorème 25.13 donne la somme directe

$$\mathcal{H} = F^\perp \oplus F^{\perp\perp}. \tag{25.59}$$

Mais  $\bar{F}$  étant également fermé nous avons la somme directe

$$\mathcal{H} = \bar{F} \oplus (\bar{F})^\perp. \tag{25.60}$$

Montrons que  $(\bar{F})^\perp = F^\perp$ . En effet si  $x \in F^\perp$  et si  $y \in \bar{F}$ , alors il existe une suite  $y_n$  dans  $F$  qui converge vers  $y$ . Pour chaque  $n$  nous avons  $\langle x, y_n \rangle = 0$  et donc  $\langle x, y \rangle = 0$  par continuité du produit scalaire.

Nous avons donc

$$\mathcal{H} = F^\perp \oplus F^{\perp\perp} = \bar{F} \oplus F^\perp. \tag{25.61}$$

Mais  $\bar{F} \subset F^{\perp\perp}$  (prendre une suite). Les espaces  $\bar{F}$  et  $F^{\perp\perp}$  étant tous deux des supplémentaires de  $F^\perp$ , nous déduisons qu'ils doivent être égaux. □

PropqiWonByiBmc

**Proposition 25.15.**

Soit  $\mathcal{H}$  un espace de Hilbert et  $F$  un sous-espace vectoriel de  $\mathcal{H}$ . Alors  $F$  est dense si et seulement si  $F^\perp = \{0\}$ .

*Démonstration.* Nous savons que

$$\mathcal{H} = F^\perp \oplus \bar{F}. \quad (25.62)$$

Donc nous avons  $\mathcal{H} = \bar{F}$  si et seulement si  $F^\perp = \{0\}$ .  $\square$

Pour vérifier si un sous-espace vectoriel d'un espace de Hilbert est dense, il suffit donc de montrer que son orthogonal est réduit à zéro.

**25.3.2 Dual, théorème de représentation de Riesz**

LemjYVcHE

**Lemme 25.16.**

*L'application linéaire*

$$\begin{aligned} \Phi: \mathcal{H} &\rightarrow \mathcal{L}(\mathcal{H}, \mathbb{K}) \\ y &\mapsto \Phi_y \end{aligned} \quad (25.63)$$

est une isométrie : nous avons  $\|\Phi_y\| = \|y\|$ .

De plus pour chaque  $y$ , l'application  $\Phi_y$  est continue.

*Démonstration.* En utilisant la définition de la norme opérateur et l'inégalité de Cauchy-Schwarz,

$$\|\Phi_y\| = \sup_{\|x\|=1} \|\Phi_y(x)\| = \sup_{\|x\|=1} |\langle x, y \rangle| \leq \sup_{\|x\|=1} \|x\| \|y\| = \|y\|. \quad (25.64)$$

Par conséquent  $\|\Phi_y\| \leq \|y\|$ . Mais d'autre part le fait que  $\Phi_y(y) = \|y\|^2$  montre que  $\|\Phi_y\| \geq \|y\|$ .

En ce qui concerne la continuité de  $\Phi_y$ , elle est garantie par le fait que c'est une application linéaire bornée via la proposition 10.81.  $\square$

**Définition 25.17.**

Le **dual** de l'espace de Hilbert  $\mathcal{H}$  est l'ensemble

$$\mathcal{H}' = \{f: \mathcal{H} \rightarrow \mathbb{K} \text{ linéaire et continue}\}. \quad (25.65)$$

Notons que dans le contexte des espaces de Hilbert nous demandons la continuité des éléments du dual parce qu'elle n'est pas automatique par la linéarité dans les cas de dimension infinie. En principe nous devrions préciser dual *topologique*, mais nous ne le ferons pas systématiquement lorsque le contexte parle clairement de topologie (ce qui est le cas lorsqu'on parle d'espaces de Hilbert). De temps en temps le dual *algébrique* d'un espace est noté  $E^*$ ; dans ce cas la continuité n'est pas demandée.

ThoQgTovL

**Théorème 25.18** (Théorème de représentation de Riesz, thème 21).

Soit un espace de Hilbert  $\mathcal{H}$  sur le corps  $\mathbb{K}$  ( $\mathbb{R}$  ou  $\mathbb{C}$ ). L'application

$$\begin{aligned} \Phi: \mathcal{H} &\rightarrow \mathcal{H}' \\ y &\mapsto \Phi_y \end{aligned} \quad (25.66)$$

est une bijection isométrique.

*Démonstration.* Nous savons du lemme 25.16 que  $\Phi$  est une isométrie. Nous devons seulement montrer que  $\Phi$  est surjective. L'application nulle est dans l'image de  $\Phi$ . Soit  $f \in \mathcal{H}'$  non nulle. Par continuité nous savons que  $F = \ker(f)$  est fermé, donc

$$\mathcal{H} = \ker(f) \oplus (\ker f)^\perp \quad (25.67)$$

EQooFUCWooUjYcZG

par le théorème 25.13.



(1) **Une base adaptée**

Nous considérons une base de  $\mathcal{H}$  adaptée à la décomposition 25.67, c'est-à-dire

- $(u_s)_{s \in S}$  une base de  $\ker f$ ,
- $(v_t)_{t \in T}$  une base de  $(\ker f)^\perp$ .

(2)  **$(\ker f)^\perp \cap \ker(f) = \{0\}$** 

Si  $v \in (\ker f)^\perp$  alors pour tout  $w \in \ker(f)$  nous avons  $\langle v, w \rangle = 0$ . En particulier si  $v$  lui-même est dans  $\ker(f)$  alors  $\langle v, v \rangle = 0$  et  $v = 0$ .

(3) **Une base encore plus adaptée** Si  $v$  est non nul dans  $(\ker f)^\perp$ , alors  $f(v) \neq 0$  (sinon il serait dans l'intersection entre  $\ker(f)$  et  $(\ker f)^\perp$ ). Mais  $f(v) \neq 0$  dans  $\mathbb{K}$  implique que  $f(\mathbb{K}v) = \mathbb{K}$ , c'est-à-dire que  $\{f(v)\}$  est une base de  $\text{Image}(f)$  (l'image de  $f$  est  $\mathbb{K}$ ).

Le théorème du rang (théorème 4.49) assure alors que

$$\{v\} \cup \{u_s\}_{s \in S} \quad (25.68)$$

est une base de  $\mathcal{H}$  avec  $v \in (\ker f)^\perp$  et  $u_s \in \ker f$ . Nous choisissons  $v$  pour avoir  $\|v\| = 1$ .

(4) **Existence**

Nous pouvons maintenant prouver l'existence de  $y$  tel que  $\Phi_y = f$ . Prouvons qu'en posant  $y = \overline{f(v)}v$  nous avons  $\Phi_y = f$ . Pour cela, un peu de calcul :

$$\Phi_y(v) = \langle v, y \rangle = f(v)\langle v, v \rangle = f(v) \quad (25.69a)$$

$$\Phi_y(u_s) = \langle u_s, y \rangle = 0. \quad (25.69b)$$

Par conséquent  $\Phi_y$  et  $f$  coïncident sur une base de  $\mathcal{H}$ .

(5) **Unicité**

En ce qui concerne l'unicité, d'abord si  $\Phi_y = f$  alors nous devons avoir  $y \in (\ker f)^\perp$  et par conséquent  $y = \lambda v$  pour un certain  $\lambda \in \mathbb{K}$ . Nous avons alors

$$\Phi_y(v) = \bar{\lambda}\langle v, v \rangle = \bar{\lambda}. \quad (25.70)$$

Pour que cela soit égal à  $f(v)$ , nous fixons  $\lambda = \overline{f(v)}$ .

□

Notons que nous avons réellement utilisé le théorème du rang pour l'unicité. Si nous ne demandions pas l'unicité, alors nous n'avions pas besoin du fait que  $\dim(\ker f)^\perp = 1$ , et nous aurons donc pu parler de formes plus générales à valeurs dans  $\mathbb{K}^n$ .

## 25.3.3 Séparabilité

DEFooSFOJooGICSbT

**Définition 25.19.**

Un espace topologique est **séparable** si il possède une partie dénombrable dense.

DEFooQVPHooJaSWyF

**Définition 25.20 ([687]).**

Si  $E$  est un espace vectoriel normé nous disons que  $\Delta$  est une partie **totale** de  $E$  si  $\text{Span}(\Delta)$  est dense dans  $E$ . Attention : nous rappelons que  $\text{Span}(\Delta)$  est l'ensemble des combinaisons linéaires finies d'éléments de  $\Delta$ .

PROPooZMWHooVwvNBY

**Proposition 25.21.**

Un espace vectoriel normé est séparable si et seulement si il possède une partie totale dénombrable.

*Démonstration.* Si  $\Delta$  est une partie dénombrable de  $E$ , alors le  $\mathbb{Q}$ -espace vectoriel  $\text{Span}_{\mathbb{Q}} \Delta$  est dénombrable, et sa fermeture est la même que celle de  $\text{Span}_{\mathbb{K}} \Delta$ . □

**Définition 25.22.**

Une famille  $(u_i)_{i \in I}$  d'éléments de  $\mathcal{H}$  indexée par un ensemble quelconque  $I$  est un **système orthonormé** si

- (1)  $\|u_i\| = 1$  pour tout  $i \in I$ ,  
 (2)  $u_i \perp u_j$  pour tout  $i \neq j$ .

Notons que si  $(u_n)_{n \in \mathbb{N}}$  est un système orthonormé dénombrable alors en utilisant les formules de la proposition 25.4, nous avons

$$\left\| \sum_{k=1}^n \xi_k u_k \right\|^2 = \sum_{k=1}^n |\xi_k|^2. \quad \text{EqCLQbMy} \quad (25.71)$$

**Exemple 25.23.**

Dans l'ensemble  $L^2(a, b)$  avec  $b - a = L$  l'ensemble des fonctions

$$e_n(t) = \frac{1}{\sqrt{b-a}} \exp\left(\frac{2\pi}{L} int\right) \quad \text{EqNXguH} \quad (25.72)$$

avec  $n \in \mathbb{Z}$  forme un système orthonormé. En effet

$$\langle e_n, e_n \rangle = \frac{1}{b-a} \int_a^b \exp\left(\frac{2\pi}{L} int\right) \exp\left(-\frac{2\pi}{L} int\right) dt = 1 \quad (25.73)$$

et

$$\langle e_n, e_m \rangle = \frac{1}{b-a} \int_a^b e^{\frac{2\pi}{L}(n-m)t} dt = 0. \quad (25.74)$$

Dans cette intégrale nous utilisons le fait que  $b = a + (b-a)$  pour simplifier les expressions en cours de calcul.

La famille (25.72) est le **système trigonométrique** de  $L^2(a, b)$ . On en parle aussi dans [647].

△

PROPooMOQRooCPFnPC

**Proposition 25.24.**

*Une partie orthonormée est libre.*

*Démonstration.* Soit  $(u_i)_{i \in I}$  une famille orthonormée de  $\mathcal{H}$  et une combinaison linéaire *finie* nulle :

$$\sum_{k=1}^n a_k u_k = 0. \quad (25.75)$$

Nous développons la somme en utilisant les formules de la proposition 25.4 :

$$0 = \left\| \sum_k a_k u_k \right\|^2 = \sum_k |a_k|^2 \|u_k\|^2 + 2 \sum_{k < l} \langle a_k u_k, a_l u_l \rangle. \quad (25.76)$$

La famille étant orthonormée, les choses se simplifient en

$$\sum_k |a_k|^2 = 0, \quad (25.77)$$

ce qui signifie que  $a_k = 0$  pour tout  $k$ . □

Avant de continuer nous devons définir comment nous calculons des sommes sur des ensembles quelconques. Si  $I$  est un ensemble et si pour chaque  $i \in I$  nous avons un nombre réel positif  $a_i$ , alors nous définissons

$$\sum_{i \in I} a_i = \sup_{\substack{J \subset I \\ \text{fini}}} \sum_{j \in J} a_j. \quad (25.78)$$

Cela est discuté dans la section 10.8.

PropHKqVHj

**Proposition 25.25** (Inégalités de Bessel).

*Soit  $\mathcal{H}$  un préhilbert. Si  $(u_i)_{i \in I}$  est un système orthonormé et si  $x \in \mathcal{H}$ , alors*

$$\sum_{i \in I} |\langle x, u_i \rangle|^2 \leq \|x\|^2. \quad \text{EqQooWCYZooQFzGaw} \quad (25.79)$$

*Démonstration.* Les éléments de la somme étant des réels positifs, la proposition 10.122 fonctionne pour décrire les sommes.

Posons  $c_i(x) = \langle x, u_i \rangle$ . Pour toute partie finie  $J \subset I$  nous avons

$$0 \leq \|x - \sum_{j \in J} c_j(x) u_j\|^2 = \|x\|^2 - 2 \operatorname{Re} \sum_j \langle x, c_j(x) u_j \rangle + \sum_j |c_j(x)|^2. \quad (25.80)$$

Mais en tenant compte du fait que

$$\langle x, c_j(x) u_j \rangle = \overline{c_j(x)} \langle x, u_j \rangle = |c_j(x)|^2, \quad (25.81)$$

nous restons avec

$$\|x - \sum_{j \in J} c_j(x) u_j\|^2 = \|x\|^2 - \sum_{j \in J} |c_j(x)|^2. \quad (25.82)$$

Finalement,

$$\sum_{j \in J} |c_j(x)|^2 \leq \|x\|^2. \quad (25.83)$$

Ayant cette inégalité pour toute partie finie de  $I$ , nous l'avons encore pour le supremum.  $\square$

### Proposition 25.26.

Soit  $\mathcal{H}$  un préhilbert et une famille orthonormée  $(u_i)_{i \in I}$ . Si

$$x = \sum_{i \in I} \xi_i u_i \quad (25.84)$$

alors  $\xi_i = \langle x, u_i \rangle$ .

*Démonstration.* Nous appliquons l'application  $\Phi_{u_k}$  du théorème de représentation de Riesz<sup>9</sup> à l'équation (25.84).

Si  $I$  est dénombrable, alors permuter  $\Phi_{u_k}$  avec la somme consiste à invoquer la continuité, et permuter la limite des sommes partielles avec  $\Phi_{u_k}$  (l'application  $\Phi_{u_k}$  est continue parce qu'isométrique).

Sinon, il faut utiliser la proposition 10.133. Il faut donc montrer que la famille  $\Phi_{u_k}(\xi_i u_i)$  est sommable. Cela est fort vrai parce que cette famille ne contient en réalité qu'un seul élément non nul, celui avec  $i = k$ , qui vaut  $\xi_k$ . Au final nous avons :

$$\langle x, u_k \rangle = \sum_{i \in I} \Phi_{u_k}(\xi_i u_i) = \xi_k. \quad (25.85)$$

$\square$

### 25.3.4 Base hilbertienne

#### Définition 25.27.

Une **base orthonormée** est une famille dénombrable orthonormée et totale<sup>10</sup>. Cela sera souvent aussi appelé une **base hilbertienne**.

#### 25.28.

Cette définition demande quelques remarques.

- (1) La notion de base hilbertienne n'est pas la même notion de base qu'en algèbre. En effet pour avoir une base algébrique d'un espace vectoriel, nous demandons que les éléments soient des combinaisons linéaires *finies* des éléments de la base, tandis qu'ici en demandant que la partie soit totale nous demandons simplement que les combinaisons linéaires finies soient denses.
- (2) Nous allons voir qu'un espace de Hilbert est généré par les sommes *infinies* de vecteurs d'une base hilbertienne avec des coefficients qui forment une suite dans  $\ell^2$ .

9. Théorème 25.18.

10. Définition 25.20

- (3) Nous ne demandons pas que la famille soit libre ? La belle affaire. Une famille orthogonale est toujours libre, proposition 25.24.
- (4) Nous verrons dans la proposition 25.37 que toute base hilbertienne est maximale (définition 25.33), c'est-à-dire que si  $\langle x, b \rangle = 0$  pour tout  $b$  dans la base, alors  $x = 0$ .

LEMooHWOBooQJKdTD

**Lemme 25.29** ([688]).

Si  $\mathcal{E}$  est une base hilbertienne<sup>11</sup> de l'espace de Hilbert  $\mathcal{H}$  et si  $x \in \mathcal{H}$  alors l'ensemble

$$\{e \in \mathcal{E} \text{ tel que } \langle x, e \rangle \neq 0\} \quad (25.86)$$

est au plus dénombrable.

*Démonstration.* Notons qu'ici,  $\mathcal{H}$  n'est pas supposé séparable. Nous savons par l'inégalité de Bessel (25.79) que

$$\sum_{e \in \mathcal{E}} |\langle x, e \rangle|^2 \leq \|x\|^2. \quad (25.87)$$

Donc si  $\epsilon > 0$  est donné, l'ensemble  $\{e \in \mathcal{E} \text{ tel que } |\langle x, e \rangle| > \epsilon\}$  est fini. Or

$$\{e \in \mathcal{E} \text{ tel que } \langle x, e \rangle \neq 0\} = \{e \in \mathcal{E} \text{ tel que } |\langle x, e \rangle| > 0\} = \bigcup_{n=1}^{\infty} \{e \in \mathcal{E} \text{ tel que } |\langle x, e \rangle| > \frac{1}{n}\}. \quad (25.88)$$

Bref, cet ensemble est une union dénombrable d'ensembles finis. Il est donc dénombrable.  $\square$

**Remarque 25.30.**

Le lemme 25.29 ne signifie pas que la base  $\mathcal{E}$  doive être dénombrable. Il signifie seulement que pour chaque  $x$  séparément, seule une partie dénombrable de  $\mathcal{E}$  est nécessaire.

CORooFROTooNupAQs

**Corolaire 25.31** ([688]).

Si un espace de Hilbert possède une base hilbertienne dénombrable, alors toutes ses bases hilbertiennes sont dénombrables.

*Démonstration.* Soient  $\mathcal{E}$  et  $\mathcal{F}$  des bases hilbertiennes de l'espace de Hilbert  $\mathcal{H}$ , en supposant que  $\mathcal{E}$  soit dénombrable. Pour chaque  $f \in \mathcal{F}$  nous avons  $\langle f, e \rangle \neq 0$  pour au moins un  $e \in \mathcal{E}$ , sinon en vertu de la décomposition (25.84) de  $f$  dans la base  $\mathcal{E}$ , nous aurions  $f = 0$ . Nous avons donc

$$\mathcal{F} \subset \bigcup_{e \in \mathcal{E}} \{f \in \mathcal{F} \text{ tel que } \langle f, e \rangle \neq 0\}, \quad (25.89)$$

EQooKRTFooNABXoH

alors que le lemme 25.29 indique que chacun des ensembles de l'union est au plus dénombrable. La partie  $\mathcal{F}$  est donc une union dénombrable d'ensemble dénombrables. Elle est dénombrable.  $\square$

**Remarque 25.32.**

En travaillant un peu plus sur la notion de cardinalité, le corolaire 25.31 indique que toutes les bases hilbertiennes ont même cardinalité. En effet en laissant tomber l'hypothèse de dénombrabilité sur  $\mathcal{E}$ , l'inclusion (25.89) donne que  $\mathcal{F}$  est une union de  $\text{Card}(\mathcal{E})$  ensembles dénombrables et est alors de cardinalité  $\text{Card}(\mathcal{E})$ .

DEFooRFATooDRKWoj

**Définition 25.33.**

Une partie orthonormale  $B$  est **maximale** si le seul  $x \in \mathcal{H}$  vérifiant  $\langle x, b \rangle = 0$  pour tout  $b \in B$  est  $x = 0$ .

LEMooXIECooCAQeJN

**Lemme 25.34** ([1, 689]).

Tout espace de Hilbert possède une partie orthonormale maximale.

11. Définition 25.27.

*Démonstration.* Soit  $\mathcal{H}$  un espace de Hilbert et  $\mathcal{O}$ , l'ensemble de parties orthonormales de  $\mathcal{H}$ , ordonné<sup>12</sup> par l'inclusion. Cet ensemble est inductif. En effet soit une partie totalement ordonnée  $\{A_i\}_{i \in I}$  de  $\mathcal{O}$  : chaque  $A_i$  est une partie orthonormale de  $\mathcal{H}$ , et de plus pour  $i, j \in I$  nous avons soit  $A_i \subset A_j$  soit  $A_j \subset A_i$ . Nous pouvons considérer

$$A = \bigcup_{i \in I} A_i. \quad (25.90)$$

Cela est encore une partie orthonormée de  $\mathcal{H}$  parce que si  $x, y \in A$ , alors il existe  $i, j \in I$  tels que  $x \in A_i$  et  $y \in A_j$ . Vu que  $\{A_i\}_{i \in I}$  est totalement ordonné nous supposons pour fixer les idées que  $A_i \subset A_j$ . Alors  $x$  et  $y$  sont dans  $A_j$  qui est une partie orthonormée ; nous en déduisons que  $\langle x, y \rangle = 0$  et donc que  $A$  est une partie orthonormée. C'est-à-dire :  $A \in \mathcal{O}$ . Par ailleurs,  $A$  est un majorant de  $\{A_i\}_{i \in I}$  pour l'inclusion parce que  $A_i \subset A$  pour tout  $i$ .

Nous avons prouvé que  $\mathcal{O}$  est un ensemble inductif. Le lemme de Zorn 1.37 nous dit alors que  $\mathcal{O}$  possède un élément maximum. Cet élément est une partie orthonormale incluse dans aucune autre partie orthonormale. C'est-à-dire qu'il est une partie orthonormale maximale au sens de la définition 25.33.  $\square$

PROPooENTTooIplRAS

**Proposition 25.35** ([689]).

*Si un espace de Hilbert possède une partie orthonormale maximale dénombrable, alors toutes les parties orthonormales sont dénombrables (ou finies).*

*Démonstration.* Soit  $B$  une partie orthonormale maximale, et  $A$  une partie orthonormale. Pour chaque  $b \in B$  nous notons

$$A(b) = \{a \in A \text{ tel que } \langle a, b \rangle \neq 0\} \cup \{0\}. \quad (25.91)$$

Un élément de  $A$  qui ne serait dans aucun des  $A(b)$  serait perpendiculaire à tous les éléments de  $B$ , et serait donc l'élément nul. Mais l'élément nul est dans tous les  $A(b)$  ; donc nous avons

$$A \subset \bigcup_{b \in B} A(b). \quad (25.92)$$

Or par l'inégalité de Bessel, l'ensemble  $A(b)$  est dénombrable. Par conséquent  $A$  est inclus dans une union dénombrable d'ensembles dénombrables ;  $A$  est dénombrable.  $\square$

Le fait que tout espace de Hilbert possède une base hilbertienne est vrai. Nous allons démontrer ce résultat d'abord pour les espaces séparables et ensuite, indépendamment, en général. Si vous vous la sentez de maîtriser la proposition 25.37, vous pouvez sauter la 25.36.

PROPooDRFBooWTcunC

**Proposition 25.36.**

*Tout espace de Hilbert séparable possède une base hilbertienne.*

*Démonstration.* Vu que nous supposons avoir un espace de Hilbert séparable, il possède une partie totale dénombrable par la proposition 25.21. Soit  $(v_n)_{n \in \mathbb{N}}$  une telle partie. Quitte à supprimer les  $v_i$  qui sont combinaisons linéaires des précédents, nous pouvons supposer que cette partie est libre. Nous considérons l'espace vectoriel

$$F_n = \text{Span}\{v_1, \dots, v_n\}. \quad (25.93)$$

Sur  $F_n$  nous pouvons appliquer un procédé de Gram-Schmidt pour construire une base orthonormée  $\{u_1, \dots, u_n\}$  de  $F_n$  au sens usuel. En considérant  $F_{n+1}$  et en recommençant, les vecteurs  $u_1, \dots, u_n$  ne changent pas, mais nous obtenons un vecteur  $u_{n+1}$ .

Nous construisons ainsi une suite  $(u_n)$  qui est alors orthonormée au sens des espaces de Hilbert. Nous devons encore prouver qu'il s'agit d'un ensemble total. Cela est simplement dû au fait que tout élément de  $\text{Span}\{v_n\}$  est contenu dans  $\text{Span}\{u_n\}$  parce que  $\text{Span}$  ne considère que des combinaisons linéaires finies.  $\square$

---

12. Définition 1.23.

PROPooLDXFooRaxBsI

**Proposition 25.37** ([689]).

À propos de parties orthonormales maximales.

ITEMooVUFXooDrVwum

(1) Une partie d'un espace de Hilbert est orthonormale maximale<sup>13</sup> si et seulement si elle est une base hilbertienne.

ITEMooZFENooQnSlrv

(2) Tout espace de Hilbert possède une base hilbertienne.

*Démonstration.* Soit une partie orthonormale maximale  $B$  et la fermeture de son espace engendré :  $F = \overline{\text{Span}(B)}$ . Pour que  $B$  soit une base, nous devons démontrer que  $F = \mathcal{H}$ . Pour cela nous considérons  $x \in \mathcal{H}$  et nous utilisons le théorème de projection orthogonale 25.9 pour mentionner le fait que

$$x - \text{proj}_F(x) \perp F \quad (25.94)$$

Cela dit que  $x - \text{proj}_F(x)$  est un vecteur orthogonal en particulier à tous les éléments de  $B$  ; par maximalité nous avons  $x - \text{proj}_F(x) = 0$ , c'est-à-dire  $x = \text{proj}_F(x)$  ou encore  $x \in F$ . Cela prouve que  $F = \mathcal{H}$ .

Note : pour être pointilleux, nous aurions dû travailler non avec  $x - \text{proj}_F(x)$ , mais avec le vecteur normalisé à 1.

En ce qui concerne le second point, nous invoquons le lemme 25.34 pour dire que tout espace de Hilbert possède une partie orthonormale maximale. Ensuite la première partie de cette proposition nous dit que cette dernière est une base hilbertienne.  $\square$

PROPooEYRSooCRPjdK

**Proposition 25.38** (Unicité de la décomposition dans une base hilbertienne[1]).

Soit un espace de Hilbert  $\mathcal{H}$  muni d'une base hilbertienne  $\{u_i\}_{i \in I}$ . Alors :

(1) Si

$$\sum_{i \in I} a_i u_i = 0, \quad (25.95)$$

alors  $a_i = 0$  pour tout  $i \in I$ .

(2) Si  $\sum_{i \in I} a_i u_i = \sum_{i \in I} b_i u_i$ , alors  $a_i = b_i$  pour tout  $i$  dans  $I$ .

*Démonstration.* Supposons que  $\sum_i a_i u_i = 0$ . La proposition 25.26 nous dit alors que

$$a_j = \left\langle \sum_i a_i u_i, u_j \right\rangle = \langle 0, u_j \rangle = 0. \quad (25.96)$$

Pour la même raison, si

$$x = \sum_{i \in I} a_i u_i = \sum_{i \in I} b_i u_i, \quad (25.97)$$

alors, pour chaque  $i \in I$ , nous avons  $a_i = \langle x, u_i \rangle$  et  $b_i = \langle x, u_i \rangle$ , c'est-à-dire  $a_i = b_i$ .  $\square$

Voici un petit résumé de ce que nous avons vu en termes de dénombrabilité et séparabilité.

THOooMKNFooVrCNGA

**Théorème 25.39.**

Pour un espace de Hilbert, les choses suivantes sont équivalentes.

ITEMooSJKVooFIibwg

(1) L'espace est séparable<sup>14</sup>.

ITEMooQIZLooYdtYqF

(2) L'espace possède au moins une base hilbertienne<sup>15</sup> dénombrable.

ITEMooHYSXooOubwUy

(3) Toutes les bases hilbertiennes sont dénombrables.

ITEMooMZICooNBAVum

(4) Toute partie libre est dénombrable.

*Démonstration.* Plein de résultats à citer ...

13. Définition 25.33.

14. Définition 25.19.

15. Définition 25.27.

(1) **(1) implique (2)**

est la proposition 25.36.

(2) **(2) implique (3)**

est la proposition 25.31.

(3) **(3) implique (4)**

Le procédé de Gram-Schmidt met en bijection une partie libre avec une partie orthonormale (qui engendre le même espace, mais c'est une autre affaire). Or lorsque l'espace de Hilbert possède une base dénombrable, toutes les parties orthonormales sont dénombrables par la proposition 25.35.

(4) **(4) implique (1)** La proposition 25.37(2) dit que tout espace de Hilbert possède des bases hilbertiennes. Une telle base est forcément une partie libre parce que toute famille orthonormale est libre (proposition 25.24), et donc dénombrable par hypothèse. À ce point nous avons montré que notre espace de Hilbert possédait une base hilbertienne dénombrable. Cela implique qu'il est séparable par la proposition 25.21.

□

RemfdJcQF

**Remarque 25.40.**

À mon avis il doit exister un théorème de complétion de base hilbertienne disant que si on a une famille orthonormée, alors elle se prolonge en base. Utilisant cela, nous trouvons une nouvelle démonstration de la proposition 25.25 en disant que la somme sur la « partie de base » est plus petite que la somme sur la « base complète ».

**25.41.**

Vu que nous n'avons l'intention de ne travailler qu'avec des espaces de Hilbert séparables et que toutes leurs bases sont dénombrables, nous n'allons travailler qu'avec des bases dénombrables, et donc des systèmes orthonormés dénombrables. Nous allons conventionnellement les indiquer par  $\mathbb{N}$ .

La proposition suivante explique que la notion de projection est compatible avec la décomposition d'un vecteur dans un système orthonormé.

**Proposition 25.42.**

Soit  $(u_k)_{k \geq 1}$  un système orthonormé d'un préhilbert  $\mathcal{H}$ . Soient

$$x = \sum_{k=1}^{\infty} \xi_k u_k \quad (25.98)$$

et

$$F = \text{Span}\{u_1, \dots, u_n\}. \quad (25.99)$$

Alors

$$\text{proj}_F(x) = \sum_{k=1}^n \xi_k u_k. \quad (25.100)$$

*Démonstration.* Nous allons dans un premier temps montrer que

$$y = x - \sum_{k=1}^n \xi_k u_k \quad (25.101)$$

est dans  $F^\perp$ . Pour cela nous calculons

$$\langle x - \sum_{k=1}^n \xi_k u_k, u_j \rangle = \langle \sum_{k=n+1}^{\infty} \xi_k u_k, u_j \rangle = 0 \quad (25.102)$$

où nous avons utilisé la continuité du produit scalaire pour permuter la somme (infinie) et le produit. Étant donné que  $y \in F^\perp$  nous avons  $\text{proj}_F y = 0$  par le point (2) du théorème 25.13.

D'autre part  $\text{proj}_F y$  peut être calculé selon

$$\text{proj}_F y = \text{proj}_F x - \sum_{k=1}^n \xi_k \text{proj}_F u_k \quad (25.103)$$

tandis que  $\text{proj}_F u_k = u_k$  lorsque  $1 \leq k \leq n$ . Par conséquent l'annulation de  $\text{proj}_F y$  donne

$$\text{proj}_F x = \sum_{k=1}^n \xi_k u_k, \quad (25.104)$$

donc le résultat. □

ThooRArDp

**Théorème 25.43** (Meilleure approximation).

Soit  $\{u_i\}_{i \in I}$  une famille orthonormée de l'espace de Hilbert  $\mathcal{H}$ . Alors pour tout  $x \in \mathcal{H}$  et pour tout  $J$  fini dans  $I$  et pour toute famille de nombres complexes  $(a_j)_{j \in J}$  nous avons

$$\left\| \sum_{j \in J} \langle x, u_j \rangle u_j - x \right\| \leq \left\| \sum_{j \in J} a_j u_j - x \right\|. \quad (25.105)$$

Ce théorème exprime le fait que les nombres  $\langle x, u_i \rangle$  sont les meilleurs coefficients à mettre devant les  $u_i$  pour approximer  $x$ .

**Corolaire 25.44.**

Soit  $\{u_i\}_{i \in I}$  une famille orthonormée de  $\mathcal{H}$ . Pour tout  $x \in \mathcal{H}$  et pour toutes parties finies  $J, K$  de  $I$  avec  $J \subset K$  nous avons

$$\left\| \sum_{j \in K} \langle x, u_j \rangle u_j - x \right\| \leq \left\| \sum_{j \in J} \langle x, u_j \rangle u_j - x \right\|. \quad (25.106)$$

Ce corolaire exprime le fait que plus on prend de termes de la forme  $\langle x, u_i \rangle u_i$ , mieux c'est.

PropzaKXHq

**Proposition 25.45.**

Soit  $\mathcal{H}$  un espace de Hilbert et  $(u_n)$  un système orthonormé dans  $\mathcal{H}$ . Si  $(\xi_n)_{n \in \mathbb{N}}$  est une suite dans  $\ell^2$  alors la série

$$\sum_{n=1}^{\infty} \xi_n u_n \quad (25.107)$$

converge dans  $\mathcal{H}$ .

Autrement dit l'application

$$\begin{aligned} S: \mathcal{H} &\rightarrow \ell^2 \\ x &\mapsto (\langle x, u_n \rangle)_{n \geq 1} \end{aligned} \quad (25.108)$$

est surjective.

*Démonstration.* Nous allons montrer que la série  $\sum_{n=1}^{\infty} \xi_n u_n$  est de Cauchy, c'est-à-dire que la limite

$$\lim_{n \rightarrow \infty} \left\| \sum_{k=n}^{n+p} \xi_k u_k \right\| = 0 \quad (25.109)$$

est uniforme en  $p$ . Cela est un corolaire de la formule (25.71) parce que

$$\left\| \sum_{k=n}^{n+p} \xi_k u_k \right\|^2 = \sum_{k=n}^{n+p} |\xi_k|^2. \quad (25.110)$$

Mais si  $(\xi_n)$  est dans  $\ell^2$ , pour tout  $\epsilon$ , il existe  $N$  tel que si  $n > N$  alors le membre de droite est inférieur à  $\epsilon$  indépendamment de  $p$ . □



### 25.3.5 Décomposition dans une base hilbertienne

Étant donné une base hilbertienne de  $\mathcal{H}$ , nous notons

$$c_k(x) = \langle x, u_k \rangle. \quad (25.111)$$

Dans le théorème suivant (et d'ailleurs partout), les sommes sur  $I$  sont prises au sens de la définition 10.118.

ThoyAjoqP

**Théorème 25.46** (Décomposition dans une base orthogonale).

Soit  $\mathcal{H}$  un espace de Hilbert séparable  $\{u_i\}_{i \in I}$  une base orthonormée ( $I$  est un ensemble dénombrable quelconque).

ItemQGwoIx

(1) Pour tout  $x \in \mathcal{H}$  nous avons

$$x = \sum_{i \in I} \langle x, u_i \rangle u_i \quad (25.112)$$

EoPUPTXJ

où la somme est prise au sens de la définition 10.118. En particulier, la somme converge de façon commutative.

(2) Si  $\{e_i\}_{i \in I}$  est une famille orthonormée qui satisfait la décomposition (25.112) pour tout  $x \in \mathcal{H}$  alors  $\{e_i\}$  est une base hilbertienne.

(3) Nous avons l'identité de **Plancherel**

$$\|x\|^2 = \sum_{i \in I} |\langle x, u_i \rangle|^2. \quad (25.113)$$

Le point (5) nous indiquera que cette égalité est en fait suffisante pour dire que nous avons une base.

(4) Nous avons l'identité de **Parseval**

$$\langle x, y \rangle = \sum_{i \in I} \langle x, u_i \rangle \overline{\langle y, u_i \rangle}. \quad (25.114)$$

EoHZxjtKt

(5) Si  $\{u_i\}$  est une famille de vecteurs unitaires vérifiant l'identité de Plancherel, alors c'est une base hilbertienne.

ItemQGwoIx

(6) Si  $\{u_i\}_{i \in I}$  est une base hilbertienne, la suite  $n \mapsto |\langle x, u_n \rangle|$  appartient à  $\ell^2(I)$ .

*Démonstration.* (1) Étant donné que le système  $\{u_i\}_{i \in I}$  est total, nous pouvons considérer une suite de combinaisons linéaires finies des  $u_i$  qui converge vers  $x$ . Nous écrivons

$$x_n = \sum_{j \in J_n} a_{n,j} u_j \quad (25.115)$$

et  $x_n \rightarrow x$  dans  $\mathcal{H}$ . Les ensembles  $J_n$  sont des sous-ensembles finis de  $I$ . Nous pouvons les choisir de telle sorte que  $J_n \subset J_{n+1}$  et  $\bigcup_{n \in \mathbb{N}} J_n = I$ . Ce choix correspond à éventuellement prendre  $a_{n,j} = 0$  pour toutes les valeurs de  $j$  « en trop ».

Soit  $\epsilon > 0$  et  $N$  tel que  $\|x_n - x\| < \epsilon$  pour tout  $n \geq N$ . Nous allons montrer que pour tout  $J$  fini tel que  $J_N \subset J$  nous avons  $\|\sum_{j \in J} \langle x, u_j \rangle u_j - x\| < \epsilon$ . Étant donné que

$$\sum_{j \in J_N} \langle x, u_j \rangle u_j = \sum_{j \in J} a_j u_j \quad (25.116)$$

avec

$$a_j = \begin{cases} \langle x, u_j \rangle & \text{si } j \in J_N \\ 0 & \text{sinon,} \end{cases} \quad (25.117)$$

le théorème de meilleure approximation 25.43 nous enseigne que

$$\left\| \sum_{j \in J} \langle x, u_j \rangle u_j - x \right\| \leq \left\| \sum_{j \in J_N} \langle x, u_j \rangle u_j - x \right\| < \epsilon \quad (25.118)$$

par conséquent la somme  $\sum_{i \in I} \langle x, u_i \rangle u_i$  converge vers  $x$  au sens général, et en particulier commutativement.

Les sommes étant commutatives (en particulier  $x = \sum_{i \in I} \langle x, u_i \rangle u_i$ ), et les bases hilbertiennes étant dénombrables, nous ne perdons aucune généralité en ne considérant que des bases indexées par  $\mathbb{N}$ .

- (2) Nous devons montrer que l'ensemble des combinaisons linéaires finies est dense dans  $\mathcal{H}$ . Par hypothèse, pour tout  $\epsilon$ , il existe un ensemble fini  $J$  tel que

$$\left\| \sum_{j \in J} \langle x, u_j \rangle u_j - x \right\| < \epsilon. \quad (25.119)$$

Cela prouve la densité dont nous avons besoin.

- (3) La norme étant une fonction continue, elle commute avec les sommes infinies, de telle sorte que l'égalité de Plancherel donne

$$\|x\|^2 = \left\| \sum_{i \in I} \langle x, u_i \rangle u_i \right\|^2. \quad (25.120)$$

Le système des  $u_i$  étant orthonormé,

$$\|x\|^2 = \left\| \sum_{i \in I} \langle x, u_i \rangle u_i \right\|^2 = \sum_{i \in I} |\langle x, u_i \rangle|^2. \quad (25.121)$$

- (4) Au tour de **Parseval**. Nous commençons par prouver que la somme du membre de droite converge. En utilisant l'inégalité  $|zz'| \leq |z|^2 + |z'|^2$  (valable pour  $z, z' \in \mathbb{C}$ ) et Plancherel, nous avons

$$\begin{aligned} \sum_{i \in I} |\langle x, u_i \rangle \overline{\langle y, u_i \rangle}| &\leq \sum_{i \in I} |\langle x, u_i \rangle|^2 + |\langle y, u_i \rangle|^2 \\ &\leq \|x\|^2 + \|y\|^2. \end{aligned} \quad (25.122)$$

Nous en déduisons que la famille  $\langle x, u_i \rangle \overline{\langle y, u_i \rangle}$  est (commutativement) sommable en utilisant la proposition 10.129. Par ailleurs nous savons que

$$x = \sum_{i \in I} c_i(x) u_i \quad (25.123a)$$

$$y = \sum_{i \in I} c_i(y) u_i, \quad (25.123b)$$

et le produit scalaire étant une forme bilinéaire continue,

$$\langle x, y \rangle = \sum_{i \in I} \sum_{j \in I} \langle c_i(x) u_i, \overline{c_j(y) u_j} \rangle \quad (25.124a)$$

$$= \sum_{i \in I} \sum_{j \in J} c_i(x) \overline{c_j(y)} \delta_{ij} \quad (25.124b)$$

$$= \sum_{i \in I} c_i(x) \overline{c_i(y)}. \quad (25.124c)$$

- (5) Nous utilisons l'égalité de Plancherel avec  $x = u_j$  :

$$\|u_j\|^2 = \|u_j\|^2 + \sum_{i \in I \setminus \{j\}} |\langle u_j, u_i \rangle|^2. \quad (25.125)$$

Par conséquent  $\langle u_j, u_i \rangle = 0$  dès que  $i \neq j$ . Cela prouve que le système  $\{u_i\}_{i \in I}$  est orthonormé. Nous devons encore prouver que le système est total. Pour cela nous repartons de l'équation (25.82) que nous avons déduites dans la démonstration de l'inégalité de Bessel :

$$\left\| x - \sum_{j \in J} c_j(x) u_j \right\|^2 = \|x\|^2 - \sum_{j \in J} |c_j(x)|^2. \quad (25.126)$$

Par hypothèse le membre de droite peut être rendu aussi petit que l'on veut en prenant  $J$  grand (mais fini) dans  $I$ . Le membre de gauche indique alors que le système  $\{u_i\}_{i \in I}$  est total.

- (6) L'identité de Plancherel signifie entre autres que si  $x \in \mathcal{H}$  alors  $\sum_{i \in I} |\langle x, u_i \rangle|^2$  converge. Du coup la suite  $(\langle x, u_i \rangle)_{i \in I}$  est dans  $\ell^2(I)$ . □

**Remarque 25.47.**

Nous avons décidé d'indexer les bases hilbertiennes par  $\mathbb{N}$  ; cela est légitime parce que les sommes sont commutatives. Il ne faut cependant pas perdre de vue qu'en pratique l'ensemble naturel avec lequel on indexe une base est parfois  $\mathbb{Z}$ . Un tel cas est donné par la base trigonométrique de  $L^2$ . Indexer cette dernière par  $\mathbb{N}$  plutôt que par  $\mathbb{Z}$  serait une contorsion inutile.

**Remarque 25.48.**

L'égalité de Parseval est la raison pour laquelle les physiciens écrivent souvent

$$\text{Id} = \sum_{n=1}^{\infty} |u_n\rangle\langle u_n| \quad (25.127)$$

dans les livres de mécanique quantique par exemple. Dans certains[690], nous lisons même

$$\int_{-\infty}^{+\infty} dq |q\rangle\langle q| = \hat{I}. \quad (25.128)$$

Notons que ces personnes travaillent avec un espace de Hilbert dont la base n'est pas dénombrable. Pour dire que la physique, ça n'utilise pas des mathématiques pour rire !

**Remarque 25.49.**

Par définition une base orthonormée est donc une partie dénombrable dont l'espace vectoriel engendré est dense. Un espace de Hilbert possédant une base orthonormée est donc séparable. C'est ce fait qui nous pousse à ne considérer que des espaces de Hilbert séparables ; nous n'allons donc pas étudier ce qu'il se passerait par exemple en considérant l'espace vectoriel librement engendré par les éléments de  $\mathbb{R}$ .

**Exemple 25.50.**

L'identité de Parseval (25.114), dans le cas de l'espace des fonctions continues périodiques de période  $2\pi$  signifie qu'en posant

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(s) e^{-ins}, \quad (25.129)$$

nous avons

$$\frac{1}{2\pi} \int_0^{2\pi} |f(s)|^2 = \sum_{n=-\infty}^{\infty} |c_n(f)|^2. \quad (25.130) \quad \text{EqMtuCSfz}$$

△

CorQETwUdF

**Corolaire 25.51.**

Soit  $\mathcal{H}$  un espace de Hilbert et  $(u_n)$  une base orthonormée. L'application

$$\begin{aligned} S: \mathcal{H} &\rightarrow \ell^2 \\ x &\mapsto (\langle x, u_n \rangle)_{n \in \mathbb{N}} \end{aligned} \quad (25.131)$$

est un isomorphisme d'espaces de Hilbert.

De plus l'isomorphisme réciproque est

$$\begin{aligned} S^{-1}: \ell^2 &\rightarrow \mathcal{H} \\ (\xi_n) &\mapsto \sum_{n=1}^{\infty} \xi_n u_n. \end{aligned} \quad (25.132)$$

*Démonstration.* Nous devons prouver que l'application est bijective et qu'elle vérifie

$$S(x) \cdot S(y) = \langle x, y \rangle \quad (25.133)$$

où le point dénote le produit dans  $\ell^2$ .

Pour la surjectivité, si  $(\xi_n) \in \ell^2$  alors nous savons que la somme  $\sum_n \xi_n u_n$  converge par la proposition 25.45 et par conséquent  $(\xi_n)$  est l'image par  $S$  de ce vecteur de  $\mathcal{H}$ .

Pour l'injectivité, si  $S(x) = S(y)$  alors

$$x = \sum_n \langle x, u_n \rangle u_n = \sum_n \langle y, u_n \rangle u_n = y \quad (25.134)$$

en utilisant la décomposition (25.112).

Le fait que  $S$  soit une isométrie est contenu dans Parseval. □

PROPOVPVQIooPcEFSe

**Proposition 25.52** ([1]).

Soit un espace de Hilbert séparable  $\mathcal{H}$ , un sous-espace vectoriel fermé  $V$  et une base orthonormée  $\{b_i\}_{i \in I}$  de  $\mathcal{H}$ . En posant  $v_i = \text{proj}_V(b_i)$  alors

$$C = \overline{\text{Span}\{b_i - v_i\}} \quad (25.135)$$

est l'orthogonal de  $V$ .

*Démonstration.* Juste pour rappel, lorsque nous écrivons  $v_i = \text{proj}_V(b_i)$ , nous parlons de la projection orthogonale du théorème 25.9. Donc tous les vecteurs  $b_i - v_i$  sont dans  $V^\perp$ . En passant aux limites,  $C \subset V^\perp$ .

Soit  $x \in V^\perp$  que nous décomposons dans la base  $\{b_i\}_{i \in I}$  comme  $x = \sum_{i \in I} x_i b_i$ . Posons  $c_i = b_i - v_i = b_i - \text{proj}_V(b_i)$ . Alors d'une part

$$0 = \text{proj}_V(x) = \sum_i x_i \text{proj}_V(b_i) = \sum_i x_i (b_i - c_i). \quad (25.136)$$

Nous avons utilisé la continuité de  $\text{proj}_V$  pour permuter avec la somme. D'autre part,

$$x = \sum_i x_i b_i = \sum_i x_i (b_i - c_i + c_i) = \underbrace{\sum_i x_i (b_i - c_i)}_{=0} + \sum_i x_i c_i \in C. \quad (25.137)$$

Notons que pour la dernière appartenance, il est important de prendre la fermeture pour définir  $C$ . □

**Proposition 25.53** ([1, 691]).

Toute partie orthonormée d'un espace de Hilbert séparable se prolonge en une base hilbertienne.

*Démonstration.* Soit une partie orthonormée  $\{u_i\}_{i \in I}$  de l'espace de Hilbert  $\mathcal{H}$ . Nous mentionnons que cette partie est libre et donc, par le théorème 25.39, dénombrable<sup>16</sup>. Montrons pour commencer que la partie  $V = \overline{\text{Span}\{u_i\}_{i \in I}}$  est un sous-espace vectoriel fermé de  $\mathcal{H}$ .

- (1) **Vectoriel** Soit  $v, w \in V$  et  $\epsilon > 0$ . Il existe  $a \in \text{Span}\{u_i\}$  tel que  $\|a - v\| < \epsilon$  et  $b \in \text{Span}\{u_i\}$  tel que  $\|b - w\| < \epsilon$ . Dans ce cas,

$$\|v + w - (a + b)\| \leq 2\epsilon. \quad (25.138)$$

Cela prouve que  $v + w \in V$ . Nous procédons de même pour  $\lambda v$ .

- (2) **Fermé** Par construction.

16. Avec un peu de mauvaise foi, vous pouvez quand même dire que cela n'implique pas que  $I$  lui-même soit dénombrable, mais vous pouvez le supposer pour fixer les idées.

L'orthogonal<sup>17</sup> de  $V$  est un sous-espace vectoriel de  $\mathcal{H}$ , et nous pouvons donc en considérer une base hilbertienne  $C = \{c_\alpha\}_{\alpha \in A}$ . Nous prétendons que  $C \cup U$  est une base hilbertienne<sup>18</sup> de  $\mathcal{H}$ .

- (1)  $C \cup U$  est libre Bing! Il ne faut pas le démontrer : ça ne fait pas partie de la définition d'une base hilbertienne.
- (2)  $C \cup U$  est générateur Bang! Il ne faut pas le démontrer : ça ne fait pas partie de la définition d'une base hilbertienne.
- (3)  $C \cup U$  est orthogonal Ah, voilà quelque chose à démontrer. Nous devons vérifier que les produits sont nuls. Soient  $u, v \in U$  et  $a, b \in C$ . Nous avons :
  - $\langle u, v \rangle = 0$  par hypothèse.
  - $\langle u, a \rangle = 0$  parce que les éléments de  $C$  sont orthogonaux à  $V$  et que  $u_i \in V$ .
  - $\langle a, b \rangle = 0$  parce que  $C$  est une base hilbertienne de  $V^\perp$ .

(4)  $C \cup U$  est dénombrable

L'ensemble  $C$  est dénombrable parce que c'est une base hilbertienne. Quant à  $\{u_i\}_{i \in I}$ , nous avons déjà mentionné le fait qu'il doive être dénombrable. L'union deux parties dénombrables est dénombrable.

(5)  $C \cup U$  est total

Nous devons prouver que  $\overline{\text{Span}(C \cup U)} = \mathcal{H}$  parce qu'il y a bien la fermeture qui intervient dans la définition 25.20. Pour cela nous utilisons la proposition 25.15. Si  $x \in \mathcal{H}$  alors nous avons

$$x = \text{proj}_V(x) + (x - \text{proj}_V(x)) = \sum_{i \in I} x_i u_i + \sum_{\alpha \in A} x_\alpha c_\alpha \quad (25.139)$$

pour des coefficients  $x_i$  et  $x_\alpha$ . Notons que  $I$  et  $A$  sont deux ensembles différents. Aucun des  $x_\alpha$  n'est un des  $x_i$ , ni inversement. Supposons que  $x \in \text{Span}(C \cup U)^\perp$ ; alors

$$0 = \langle x, u_j \rangle = \sum_{i \in I} x_i \underbrace{\langle u_i, u_j \rangle}_{=\delta_{ij}} + \sum_{\alpha \in A} \underbrace{\langle c_\alpha, u_j \rangle}_{=0} = x_j. \quad (25.140)$$

Donc  $x_j = 0$ . En faisant de même avec  $0 = \langle x, c_\beta \rangle = x_\beta$  nous déduisons  $x = 0$ .

□

PROPooZQIAooUEWLtg

**Proposition 25.54** ([1]).

Si  $\mathcal{H}$  est un espace de Hilbert séparable de base hilbertienne  $\{e_i\}_{i \in I}$ , alors  $\mathcal{H}'$  est un espace de Hilbert séparable dont une base hilbertienne est donnée par les formes

$$\begin{aligned} \alpha_i: \mathcal{H} &\rightarrow \mathbb{C} \\ e_j &\mapsto \delta_{ij}. \end{aligned} \quad (25.141)$$

*Démonstration.* C'est la proposition 25.6 qui fait tout.

□

### 25.3.6 Digression sur les normes opérateurs

subsecaeSywF

Le théorème 25.46 nous indique que si  $\{u_i\}_{i \in I}$  est une base hilbertienne, alors pour tout  $x \in \mathcal{H}$  nous avons

$$\text{proj}_{u_i}(x) = \langle x, u_i \rangle u_i, \quad (25.142)$$

et donc

$$\sum_{i \in I} \text{proj}_{u_i} x = x. \quad (25.143)$$

17. Voir la définition 25.11 et la proposition 25.52.

18. Définition 25.27.

Nous ne pouvons cependant pas conclure que

$$\sum_{i \in I} \text{proj}_{u_i} = \text{Id} \quad (25.144)$$

au sens de la norme opérateur de la définition 10.63. En effet en prenant  $I = \mathbb{N}$ , l'égalité (25.144) demanderait d'avoir

$$\lim_{N \rightarrow \infty} \left\| \sum_{i=1}^N \text{proj}_{u_i} - \text{Id} \right\|_{\infty} = 0, \quad (25.145)$$

or pour tout  $N$ , le vecteur  $u_{N+1}$  réalise

$$\sum_{i=1}^N \text{proj}_{u_i}(u_{N+1}) - u_{N+1} = -u_{N+1}. \quad (25.146)$$

Par conséquent pour tout  $N$  nous avons

$$\sup_{\|x\|=1} \left\| \sum_{i=1}^N \text{proj}_{u_i} x - x \right\| \geq 1. \quad (25.147)$$

Nous ne pouvons donc pas dire

$$\sum_{n=1}^{\infty} \text{proj}_{u_i} = \text{Id} \quad (25.148)$$

au sens de la norme opérateur.

Nous avons cependant la convergence au sens faible.

**Proposition 25.55.**

Soit  $\mathcal{H}$  un espace de Hilbert et  $\{u_i\}_{i \in \mathbb{N}}$  une base hilbertienne de  $\mathcal{H}$ . Au sens de la topologie faible sur l'espace des opérateurs nous avons

$$\sum_{i=1}^{\infty} \text{proj}_{u_i} = \text{Id}. \quad (25.149)$$

*Démonstration.* Pour chaque  $N \in \mathbb{N}$  et  $x \in \mathcal{H}$ , en vertu de la décomposition (25.112) nous avons

$$\sum_{i=1}^N \text{proj}_{u_i}(x) - x = \sum_{i=N+1}^{\infty} \langle x, u_i \rangle u_i. \quad (25.150)$$

Par l'orthonormalité de la base nous avons

$$\sum_{i=N+1}^{\infty} \|\langle x, u_i \rangle u_i\| = \sum_{i=N+1}^{\infty} |\langle x, u_i \rangle|, \quad (25.151)$$

dont la limite  $N \rightarrow \infty$  est zéro étant donné que la suite  $i \mapsto |\langle x, u_i \rangle|$  est dans  $\ell^2(\mathbb{R})$  par le théorème 25.46.  $\square$

### 25.3.7 Applications linéaires et continuité

Nous avons déjà vu dans l'exemple 10.82 que la fonction

$$\begin{aligned} f: H &\rightarrow H \\ e_k &\mapsto k e_k \end{aligned} \quad (25.152)$$

n'était pas continue en zéro alors qu'elle est linéaire. Nous allons maintenant voir qu'elle est un contre-exemple à la proposition 12.321. Calculons les dérivées partielles :

$$\frac{\partial f}{\partial e_j}(x) = \frac{d}{dt} \left[ f(x + t e_j) \right]_{t=0} = \frac{d}{dt} \left[ \sum_k k(x_k + t \delta_{jk}) e_k \right]_{t=0} = j e_j. \quad (25.153)$$

où nous avons permuté la somme et la dérivée en considérant la suite de fonctions  $f_k(t) = k(x_k + tu_k)e_k$ . Donc  $\frac{\partial f}{\partial e_j}(x)$  existe et est continue sur un voisinage de  $x = 0$  (c'est même constant). Nous savons pourtant que la fonction  $f$  n'est pas différentiable en zéro parce que non continue.

L'endroit qui coince dans la preuve de la proposition 12.321 est l'introduction des « contre-terms » dans l'équation (12.872). En effet les contre-terms à ajouter seraient

$$l(x) = \frac{d}{dt} \left[ f(a + s(x - a)) \right]_{t=0} \quad (25.154)$$

qui ici serait

$$l\left(\sum_k x_k e_k\right) = \frac{d}{dt} \left[ f\left(t \sum_k x_k e_k\right) \right]_{t=0} = \sum_k x_k e_k, \quad (25.155)$$

dont la convergence est plus que douteuse.

Notons que les dérivées directionnelles n'existent pas toutes, loin s'en faut : si  $u \in H$  nous avons

$$\frac{\partial f}{\partial u}(x) = \frac{d}{dt} \left[ f(x + tu) \right]_{t=0} = \frac{d}{dt} \left[ \sum_k k(x_k + tu_k)e_k \right]_{t=0} = \sum_k k u_k e_k \quad (25.156)$$

La convergence de la dernière somme n'est pas garantie pour tout  $u$ .

### Exemple 25.56.

Soit un espace de Hilbert  $\mathcal{H}$ . L'application

$$\begin{aligned} \phi: \mathcal{H} &\rightarrow \mathbb{C} \\ e_i &\mapsto 1 \end{aligned} \quad (25.157)$$

n'est pas continue.

Nous considérons les vecteurs  $y_n = \sum_{i=1}^n e_i$  et nous posons  $x_n = y_n / \|y_n\|$ . Nous avons

$$\|y_n\|^2 = \langle y_n, y_n \rangle = \sum_{i,j}^n \langle e_i, e_j \rangle = n, \quad (25.158)$$

donc  $\|y_n\| = \sqrt{n}$ , et

$$\phi(x_n) = \frac{n}{\sqrt{n}} = \sqrt{n}. \quad (25.159)$$

L'application  $\phi$  n'est donc pas bornée et pas continue non plus.  $\triangle$

### Proposition 25.57.

Soient deux espaces de Hilbert séparables<sup>19</sup>  $\mathcal{H}_1$  et  $\mathcal{H}_2$ . Soit une base orthonormée  $\{e_i\}_{i \in I}$  où  $I$  est dénombrable. Si  $\Phi: \mathcal{H}_1 \rightarrow \mathcal{H}_2$  est une bijection linéaire continue, alors

$$\Phi\left(\sum_{i \in I} x_i e_i\right) = \sum_{i \in I} x_i \Phi(e_i). \quad (25.160)$$

*Démonstration.* Soit  $\epsilon > 0$ . Vu que  $\sum_{i \in I} x_i e_i$  est une somme convergente<sup>20</sup>, il existe une partie finie  $J$  de  $I$  telle que pour tout  $K$  fini contenant  $J$ ,

$$\left\| \sum_{j \in K} x_j e_j - x \right\| < \epsilon. \quad (25.161)$$

Vu que  $\Phi$  est une isométrie linéaire, nous avons

$$\left\| \sum_{j \in K} x_j \Phi(e_j) - \Phi(x) \right\| < \epsilon, \quad (25.162)$$

et donc bien l'égalité

$$\sum_{i \in I} x_i \Phi(e_i) = \Phi(x). \quad (25.163)$$

□

19. Donc à base dénombrable, voir 25.39.

20. Définition 10.118.

La proposition suivante donne une formule pour l'inverse du  $\Phi$  du théorème de représentation de Riesz 25.18 lorsque nous avons une base hilbertienne.

**Proposition 25.58.**

Soit un espace de Hilbert séparable  $\mathcal{H}$ . Si  $\phi \in \mathcal{H}'$ , alors

$$\Phi^{-1}(\phi) = \sum_{i \in I} \phi(e_i) e_i \quad (25.164)$$

dès que  $\{e_i\}_{i \in I}$  est une base hilbertienne de  $\mathcal{H}$ .

*Démonstration.* Nous considérons la base des  $\{\alpha_i\}_{i \in I}$  de  $\mathcal{H}'$ . Remarquez au passage que  $\alpha_i = \Phi_{e_i}$ . Nous avons  $\phi = \sum_{i \in I} \phi_i \alpha_i$  pour certains  $\phi_i \in \mathbb{C}$  (qui valent  $\phi_i = \phi(e_i)$ , ça a son importance pour la suite). En utilisant la proposition 25.57 pour permuter  $\Phi$  et la somme nous avons le calcul suivant :

$$\phi = \sum_{i \in I} \phi_i \alpha_i = \sum_{i \in I} \phi_i \Phi(e_i) = \sum_{i \in I} \Phi(\phi_i e_i) = \Phi\left(\sum_{i \in I} \phi_i e_i\right) = \Phi\left(\sum_{i \in I} \phi(e_i) e_i\right). \quad (25.165)$$

En résultat :

$$\phi = \Phi\left(\sum_{i \in I} \phi(e_i) e_i\right), \quad (25.166)$$

et donc la formule (25.164).  $\square$

## 25.4 Théorème de Kochen-Specker

Le théorème suivant est central en mécanique quantique. La démonstration provient de [692] et de Wikipédia. Nous allons démontrer complètement le théorème seulement pour les espaces de Hilbert de dimension plus grande ou égale à 4.

**Théorème 25.59** (Kochen-Specker[692]).

Soit  $\mathcal{H}$  un espace de Hilbert de dimension plus grande ou égale à 3. Une fonction  $v$  sur l'ensemble des opérateurs de  $\mathcal{H}$  ne peut pas satisfaire aux conditions suivantes :

- (1) Si  $A$  et  $B$  sont compatibles, alors  $v(A + B) = v(A) + v(B)$ ,
- (2) Si  $A$  et  $B$  sont compatibles, alors  $v(AB) = v(A)v(B)$ .

Ici nous disons que deux opérateurs sont **compatibles** lorsqu'ils possèdent une base hilbertienne commune de vecteurs propres.

*Démonstration.* Soit  $\{u_n\}_{n \in \mathbb{N}}$  une base hilbertienne de  $\mathcal{H}$ . Nous notons  $\text{proj}_i$  l'opérateur de projection sur l'espace (fermé) engendré par  $u_i$ . Ce sont des opérateurs compatibles deux à deux parce que la base  $\{u_n\}_{n \in \mathbb{N}}$  est une base commune de vecteurs propres<sup>21</sup>.

D'abord nous devons avoir  $v(\mathbb{1}) = 1$ . En effet pour tout opérateur  $A$ , nous avons

$$v(A) = v(A\mathbb{1}) = v(A)v(\mathbb{1}). \quad (25.167)$$

Pour peu que  $v(A) \neq 0$ , cela nous fait  $v(\mathbb{1}) = 1$ .

En vertu du théorème 25.46, un vecteur  $x \in \mathcal{H}$  se décompose en  $x = \sum_n \langle x, u_n \rangle u_n$ , et nous avons

$$\text{proj}_i x = \langle x, u_i \rangle u_i. \quad (25.168)$$

En effet le théorème de la projection orthogonale 25.9 nous enseigne que  $\text{proj}_i x$  serait l'unique vecteur de la forme  $\lambda u_i$  tel que  $\lambda u_i - x \perp u_i$ . Il est facile de vérifier que le vecteur proposé par (25.168) vérifie cette propriété.

Une conséquence est que

$$\left( \sum_{i=0}^{\infty} \text{proj}_i \right) (x) = x. \quad (25.169)$$

21. Pour les besoins de la physique, nous remarquons que ces opérateurs sont des opérateurs hermitiens qui commutent, mais ça ne joue pas ici.



Par conséquent, par hypothèse du théorème nous devons avoir

$$\sum_i v(\text{proj}_i) = v(\mathbb{1}) = 1. \quad (25.170)$$

Étant donné que les projections sont idempotentes,

$$v(\text{proj}_i) = v(\text{proj}_i^2) = v(\text{proj}_i)^2 \quad (25.171)$$

et donc  $v(\text{proj}_i)$  doit valoir zéro ou un. Mais la relation (25.170) donne une forte contrainte sur le choix de 0 et de 1. En effet, parmi les  $v(\text{proj}_i)$ , un et un seul doit valoir 1, les autres doivent valoir 0.

Refaisant le raisonnement pour une autre base orthonormale hilbertienne, nous trouvons que les valeurs de  $v$  sur les opérateurs de projection sur les différentes directions doivent être choisies de telle façon que tout choix de base hilbertienne orthogonale contienne exactement un 1 et le reste de zéros.

Nous voudrions maintenant insister sur un point. Le problème de déterminer de façon cohérente les valeurs 0 ou 1 pour tous les  $v(\text{proj}_i)$  revient à attacher 0 ou 1 à tous les rayons de  $\mathcal{H}$  de façon que toute base orthogonale de  $\mathcal{H}$  contienne exactement un 1. Un rayon est une direction, c'est-à-dire une classe d'équivalence  $x \sim \lambda x$ . Si nous décidons de nommer « blanc » les rayons attachés à la valeur 0 et « noirs » ceux attachés à la valeur 1, le problème se réduit à colorer la boule unité de façon compatible.

Soit  $\{u_i\}_{i \in \mathbb{N}}$  une base orthogonale de  $\mathcal{H}$  numérotée de telle sorte que  $v(u_0) = 1$  et  $v(u_k) = 0$  pour  $k \neq 0$ . Nous allons maintenant nous particulariser au cas de dimension supérieure ou égale à 4. Si  $R$  est une rotation dans le plan  $\text{Span}\{u_0, u_1, u_2, u_3\}$ , alors l'ensemble

$$\{Ru_0, Ru_1, Ru_2, Ru_3, u_k\}_{k \geq 4} \quad (25.172)$$

est encore une base orthogonale de  $\mathcal{H}$  et nous avons encore  $v(u_k) = 0$  pour  $k \geq 4$ . Par conséquent un et un seul des vecteurs  $Ru_0, Ru_1, Ru_2$  ou  $Ru_3$  est colorié en noir ; les trois autres étant blancs. Le problème est maintenant complètement réduit à la dimension 4. Note : pour réduire à la dimension 3, on procède de même, mais pour conclure, il faut travailler plus.

Nous allons construire 9 bases orthogonales de  $\mathbb{R}^4$  à partir de 18 vecteurs, chacun arrivant dans exactement deux des bases. Ils sont donnés dans le tableau suivant :

|      |      |      |      |      |      |      |      |      |
|------|------|------|------|------|------|------|------|------|
| 0 1  | 0 0  | 0 0  | 1 1  | 1 -1 | 1 -1 | -1 1 | 1 1  | 1 1  |
| 0 0  | 1 0  | 0 1  | 1 1  | 1 -1 | -1 1 | 1 1  | -1 1 | 1 -1 |
| 0 0  | 0 1  | 0 0  | 1 -1 | 1 -1 | 1 1  | 1 1  | -1 1 | 1 1  |
| 0 1  | 0 0  | 1 0  | 1 -1 | -1 1 | 1 1  | 1 -1 | 1 1  | -1 1 |
| 1 0  | 1 0  | 1 1  | 1 0  | 1 1  | 1 0  | 1 0  | 1 0  | 1 -1 |
| 1 0  | 0 1  | 0 0  | -1 0 | 0 0  | 0 -1 | 0 1  | 1 0  | 0 0  |
| 1 0  | 1 0  | 1 -1 | 0 1  | 0 0  | 0 1  | 0 1  | 0 1  | 0 0  |
| -1 0 | 0 -1 | 0 0  | 0 -1 | 1 1  | -1 0 | -1 0 | 0 -1 | 1 1  |

Chaque case de ce table représente un rayon de  $\mathbb{R}^4$  ; il y en a 18 différents, chacun écrit deux fois. Une simple vérification montre que chaque colonne est un système orthogonal. La preuve du théorème de Kochen-Specker revient à montrer que nous ne pouvons pas colorier ce tableau de façon cohérente. En effet, étant donné que chaque vecteur est écrit deux fois, le tableau doit contenir un nombre pair de cases blanches et un nombre pair de cases noires.

Par ailleurs chaque colonne étant un système orthogonal, chaque colonne contient exactement une case noire ; il y a donc exactement neuf cases noires dans le tableau, ce qui est impossible.  $\square$

## 25.5 Théorème de Lax-Milgram

### Définition 25.60.

Une forme bilinéaire  $a: V \times V \rightarrow \mathbb{R}$  sur un espace vectoriel normé  $V$  est **coercitive** si il existe  $\alpha > 0$  tel que  $a(u, u) \geq \alpha \|u\|^2$  pour tout  $u \in V$ .

**Théorème 25.61** (Lax-Milgram[693]).

THOoLLUXooHyqmVL

Soit un espace de Hilbert réel  $V$  muni de différentes choses.

- (1) L'application linéaire  $L: V \rightarrow \mathbb{R}$  qui est bornée sur  $V$ . Nous notons  $\|L\|$  sa norme.
- (2) La forme bilinéaire continue  $a$  sur  $V \times V$ .
- (3) La forme  $a$  est coercitive<sup>22</sup>.

Alors

- (1) Il existe  $M > 0$  tel que  $|a(u, v)| \leq M\|u\|\|v\|$  pour tout  $u, v \in V$ .
- (2) Le problème qui consiste à chercher  $u \in V$  tel que  $a(u, v) = L(v)$  pour tout  $v \in V$  admet une unique solution. De plus cette solution vérifie l'inégalité

$$\|u\| \leq \frac{\|L\|}{\alpha} \quad (25.173)$$

EQooUAYSooKYvQBU

*Démonstration.* Évacuons deux points faciles avant de commencer les choses sérieuses. D'abord  $L$  est continue par la proposition 10.81. Ensuite, l'existence du  $M$  est donnée par la proposition 25.2.

Maintenant nous commençons.

- (1) **Reformulation en équation linéaire** La forme  $L$  est continue et donc dans le dual  $V'$ ; le théorème de Riesz 25.18 nous donne donc  $f \in V$  tel que

$$L(v) = \langle f, v \rangle \quad (25.174)$$

pour tout  $v \in V$ . De plus si  $w \in V$  est fixé, l'application  $b_w: v \mapsto a(w, v)$  est linéaire et bornée parce que

$$\|b_w\| = \sup_{\|v\|=1} |b_w(v)| = \sup_{\|v\|=1} |a(w, v)| \leq M\|w\|\|v\| = M\|w\|. \quad (25.175)$$

Encore une fois,  $b_w$  étant continue et linéaire, elle est dans  $V'$  et Riesz nous fournit un élément  $A(w) \in V$  tel que

$$b_w(v) = \langle A(w), v \rangle \quad (25.176)$$

pour tout  $v \in V$ .

Le problème variationnel  $a(u, v) = L(v)$  est équivalent à  $\langle A(u), v \rangle = \langle f, v \rangle$ . L'ensemble des solutions de cette dernière est égal à l'ensemble des solutions de l'équation

$$A(u) = f. \quad (25.177)$$

EQooLPMPooMVuYUUX

- (2)  **$A$  est linéaire**

Soient  $\alpha, \beta \in \mathbb{R}$  et  $w, z \in V$ . Nous avons pour tout  $v \in V$  :

$$\langle A(\alpha w + \beta z), v \rangle = a(\alpha w + \beta z, v) \quad (25.178a)$$

$$= \alpha a(w, v) + \beta a(z, v) \quad (25.178b)$$

$$= \alpha \langle A(w), v \rangle + \beta \langle A(z), v \rangle \quad (25.178c)$$

$$= \langle \alpha A(w) + \beta A(z), v \rangle. \quad (25.178d)$$

Étant donné que nous avons égalité pour tout  $v \in V$  nous en déduisons que  $A(\alpha w + \beta z) = \alpha A(w) + \beta A(z)$ , ce qui signifie que  $A$  est linéaire.

- (3) **Une autre propriété de  $A$**  Nous déduisons une majoration de  $\|A(v)\|^2$  lorsque ce n'est pas nul. Pour ce faire,

$$\|A(v)\|^2 = \langle A(v), A(v) \rangle \quad (25.179a)$$

$$= a(v, A(v)) \quad (25.179b)$$

$$\leq M\|v\|\|A(v)\|. \quad (25.179c)$$

En simplifiant,  $\|A(v)\| \leq M\|v\|$ . Et donc

$$\|A(v)\|^2 \leq M^2\|v\|^2. \quad (25.180)$$

EQooBQAHOoAZBdAW

---

22. Définition 25.60.

(4) **Une contraction**

Nous allons choisir une valeur de  $\rho > 0$  telle que l'application

$$T: w \mapsto w - \rho(A(w) - f) \quad (25.181)$$

soit une contraction<sup>23</sup>. Nous avons  $T(w) - T(w') = w - w' - \rho(A(w - w'))$  et donc

$$\|T(w) - T(w')\|^2 = \|w - w'\|^2 + \rho^2 \|A(w - w')\|^2 - 2\rho \langle A(w - w'), w - w' \rangle \quad (25.182a)$$

$$= \|w - w'\|^2 + \rho^2 \|A(w - w')\|^2 - 2\rho a(w - w', w - w'). \quad (25.182b)$$

Vu que le dernier terme arrive avec un signe moins, pour majorer l'expression, il faut minorer ce terme, c'est-à-dire utiliser  $a(w - w', w - w') \geq \alpha \|w - w'\|^2$ . Et en même temps nous utilisons (25.180) pour le second terme. Au final nous pouvons factoriser  $\|w - w'\|^2$  et

$$\|T(w) - T(w')\| \leq \|w - w'\| (1 + \rho^2 M^2 - 2\rho\alpha). \quad (25.183)$$

Pour que  $T$  soit contractante, il faut  $0 < P(x) < 1$  avec  $P(x) = M^2 x^2 - 2\alpha x + 1$ . Le minimum de ce polynôme est obtenu en  $x = \frac{\alpha}{M^2}$  (la formule du  $x_{\min} = -b/2a$ ) et vaut  $1 - \frac{\alpha^2}{M^2} < 1$ . Vu que par ailleurs  $\lim_{x \rightarrow \infty} P(x) = +\infty$ , et que ce polynôme passe par au moins une valeur strictement inférieure à 1, nous savons qu'il existe un  $x$  tel que  $0 < P(x) < 1$ . En donnant à  $\rho$  cette valeur, l'application  $T$  est une contraction.

(5) **Point fixe et conclusion** L'ensemble des solutions du problème (25.177) est égal à l'ensemble des points fixes de  $T(v) = v - \rho(A(v) - f)$ .

L'application  $T: V \rightarrow V$  est contractante et  $V$  est métrique et complet. Ergo le théorème de point fixe de Picard 17.34 s'applique et il existe un unique point fixe  $u \in V$  pour l'application  $T$ . Ce point fixe est l'unique solution de notre problème initial.

(6) **La majoration**

Nous savons que pour tout  $v \in V$ , la relation  $a(u, v) = L(v)$  est vérifiée. En particulier pour  $v = u$  nous avons

$$a(u, u) = L(u). \quad (25.184)$$

D'un côté nous utilisons  $a(u, u) \geq \alpha \|u\|^2$  et de l'autre,  $L(u) \leq \|L\| \|u\|$  :

$$\alpha \|u\|^2 \leq \|L\| \|u\| \quad (25.185)$$

et donc

$$\|u\| \leq \frac{\|L\|}{\alpha}. \quad (25.186)$$

Notons que  $L(u)$  et  $a(u, u)$  sont positifs.

□

EXooTTBDooUNhBOc

**Exemple 25.62** ([694]).

La borne  $\|u\| \leq \|L\|/\alpha$  est optimale au sens où il existe des cas d'égalité. En effet nous pouvons considérer

$$a(u, v) = s \langle u, v \rangle \quad (25.187)$$

et

$$L(v) = \langle f, v \rangle \quad (25.188)$$

pour un certain  $s > 0$  et  $f \in V$ . L'application  $L$  n'est autre que  $\Phi_f$  dont nous avons abondamment parlé autour du théorème de représentation de Riesz 25.18. Le lemme 25.16 nous apprend que  $L$  est une application bornée (et donc continue) de norme  $\|L\|_{V'} = \|f\|_V$ .

L'application  $a: V \times V \rightarrow \mathbb{R}$  est continue par la proposition 10.92. Elle est coercive parce que

$$a(u, u) = s \|u\|^2. \quad (25.189)$$

---

23. Définition 17.33.

Les hypothèses du théorème de Lax-Milgram 25.61 sont réunies et il existe un unique élément  $u \in V$  tel que  $a(u, v) = L(v)$  pour tout  $v$ . Cet élément n'est autre que

$$u = \frac{1}{s}f. \quad (25.190)$$

Nous avons alors  $\|u\| = \frac{1}{s}\|f\|$ .

Mais  $s$  est la constante de coercivité de  $a$  et  $\|f\|$  est la norme de  $L$ . Cette dernière affirmation est contenue dans le fait que  $\Phi$  est une isométrie dans le lemme 25.16. En reprenant les notations du théorème, nous avons l'égalité

$$\|u\| = \frac{\|L\|}{\alpha}, \quad (25.191)$$

ce qui signifie que l'inégalité du théorème est dans un certain sens optimale.  $\triangle$

PROPOOFEOZooTNPcBJ

### Proposition 25.63.

Soit un espace de Hilbert réel  $V$  et une application  $a: V \times V \rightarrow \mathbb{R}$  coercive et continue.

L'application  $\phi: V' \rightarrow V$  qui à  $L$  fait correspondre l'unique solution  $u$  de  $a(u, v) = L(v)$  est

- (1) bien définie,
- (2) linéaire,
- (3) continue.

*Démonstration.* Le fait que  $\phi$  soit bien définie est le théorème de Lax-Milgram 25.61.

L'application  $\phi$  dont nous parlons ici est une application entre deux espaces normés. Soient  $L_1, L_2 \in V'$ ; nous définissons  $u_1 = \phi(L_1)$  et  $u_2 = \phi(L_2)$ . Nous avons

$$a(u_1, v) = L_1(v) \quad (25.192a)$$

$$a(u_2, v) = L_2(v) \quad (25.192b)$$

pour tout  $v \in V$ . Par bilinéarité de  $a$  nous avons aussi

$$a(u_1 + u_2, v) = (L_1 + L_2)(v). \quad (25.193)$$

Vu que  $L_1 + L_2$  est également bornée,  $u_1 + u_2 = \phi(L_1 + L_2)$ . Pour le même type de raisons, si  $\lambda \in \mathbb{R}$ ,

$$a(\lambda u, v) = (\lambda L)(v), \quad (25.194)$$

c'est-à-dire que  $\phi(\lambda L) = \lambda\phi(L)$ .

Donc  $\phi$  est linéaire. Pour la continuité, il suffit de s'assurer que  $\phi$  est bornée. Pour cela nous utilisons la majoration (25.173) et la définition de la norme d'une application linéaire :

$$\|\phi\| = \sup_{\|L\|=1} \|\phi(L)\| \leq \sup_{\|L\|=1} \frac{1}{\alpha} \|L\| = \frac{1}{\alpha}. \quad (25.195)$$

Donc  $\phi$  est bornée et par voie de conséquence continue (proposition 10.81).  $\square$

Un problème est bien posé au sens de Hadamard si la solution est unique et dépend continument des données; nous ne rentrerons pas plus dans les détails pour l'instant : nous en reparlerons dans la définition 33.14 mais sachez qu'il existe en fin de compte autant de Hadamard qu'il y a de contextes dans lequel le mot « problème » a un sens.

Ce que dit la proposition 25.63 est que,  $a$  étant donné, le problème qui consiste à résoudre  $a(u, v) = L(v)$  est bien posé au sens de Hadamard, au moins par rapport à  $L$ .

### Théorème 25.64 (Lax-Milgram version symétrique[693]).

Nous considérons les mêmes hypothèses que celles du théorème de Lax-Milgram, c'est-à-dire un espace de Hilbert réel  $V$  muni de différentes choses.

- (1) L'application linéaire  $L: V \rightarrow \mathbb{R}$  qui est bornée sur  $V$ . Nous notons  $C$  sa norme.

(2) La forme bilinéaire continue  $a$  sur  $V \times V$ . Nous considérons  $M > 0$  tel que  $|a(u, v)| \leq M \|u\| \|v\|$  pour tout  $u, v \in V$ .

(3) La forme  $a$  est coercitive.

Nous supposons de plus que  $a$  est symétrique.

Alors un élément  $u \in V$  est tel que  $a(u, w) = l(w)$  pour tout  $w \in V$  si et seulement si il minimise la **fonctionnelle d'énergie**

$$J(w) = \frac{1}{2}a(w, w) - l(w). \quad (25.196)$$

*Démonstration.* Nous séparons les deux sens.

(1)  $\Rightarrow$  Soit  $u$ , un élément vérifiant  $a(u, w) = l(w)$  pour tout  $w \in V$ . Soit aussi un élément quelconque  $w \in V$ , et montrons que  $J(u + w) \geq J(u)$  (tout élément de  $V$  peut être écrit sous la forme  $u + w$ ). Nous avons :

$$J(u + w) = \frac{1}{2}(a(u, u) + a(u, w) + a(w, u) + a(w, w)) - l(u) - l(w) \quad (25.197a)$$

$$= J(u) + \underbrace{a(u, w) - l(w)}_{=0} + \frac{1}{2}a(w, w) \quad (25.197b)$$

$$= J(u) + \frac{1}{2}a(w, w) \quad (25.197c)$$

Vu que  $a$  est coercive, le second terme est positif et nous avons

$$J(u + w) \geq J(u), \quad (25.198)$$

ce qu'il fallait.

(2)  $\Leftarrow$

Soit  $u \in V$ , un élément minimisant la fonctionnelle  $J$ . Nous fixons  $w \in V$  et considérons la fonction  $g: \mathbb{R} \rightarrow \mathbb{R}$  définie par

$$g(\epsilon) = J(u + \epsilon w). \quad (25.199)$$

En développant un peu et en regroupant les termes,

$$g(\epsilon) = J(u) + \epsilon(a(u, w) - l(w)) + \frac{\epsilon^2}{2}a(w, w). \quad (25.200)$$

Cela est une fonction éminemment continue et dérivable; en réalité c'est un bête polynôme de degré deux. Vu que  $u$  minimise  $J$ , pour tout  $\epsilon \neq 0$  nous avons  $g(\epsilon) \geq g(0)$  ou encore :  $\epsilon = 0$  est minimum local (et même global) de  $g$ . Le polynôme (25.200) prend son minimum en  $\epsilon = 0$  si et seulement si

$$a(u, w) - l(w) = 0. \quad (25.201)$$

Vous ne me croyez pas ? Faites  $g'(\epsilon) = 0$  ou bien reprenez la formule du  $-b/2a$  pour le sommet d'une parabole, en tenant compte que  $a(w, w) \neq 0$ . Notons qu'ici encore le fait que  $a$  soit coercive joue parce que c'est cela qui nous permet d'affirmer que la parabole a un minimum et non un maximum.

□

# Chapitre 26

## Analyse complexe

ChapI CH IooXbLccl

### 26.1 Fonctions holomorphes

La dérivée complexe est discutée à la section 12.31, et la définition d'une fonction holomorphe est 12.331.

#### 26.1.1 Équations de Cauchy-Riemann

Notons que la formule (12.917) donne un **développement limité** pour les fonctions holomorphes. Si  $f$  est holomorphe en  $z_0$  alors si  $z$  est dans un voisinage de  $z_0$ , il existe une fonction  $s: \mathbb{R} \rightarrow \mathbb{C}$  telle que  $\lim_{t \rightarrow 0} s(t)/t = 0$  et

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + s(|z - z_0|). \quad \text{EqptwBFG (26.1)}$$

**26.1** (Quelque (abus de) notations).

Nous utilisons l'application

$$\begin{aligned} \varphi: \mathbb{R}^2 &\rightarrow \mathbb{C} \\ (x, y) &\mapsto x + iy. \end{aligned} \quad (26.2)$$

Nous notons

$$\partial_x f = \partial_1(f \circ \varphi) \quad (26.3a)$$

$$\partial_y f = \partial_2(f \circ \varphi). \quad (26.3b)$$

Avec ces définitions, nous introduisons les opérateurs

$$\frac{\partial}{\partial z} = \partial = \frac{1}{2} \left( \frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \quad (26.4a)$$

$$\frac{\partial}{\partial \bar{z}} = \bar{\partial} = \frac{1}{2} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \quad \text{SUBEQooNYQKooYqLQzE (26.4b)}$$

La raison de ces définitions est d'avoir l'égalité

$$f'(z_0) = (\partial_z f)(z_0) \quad (26.5)$$

que nous verrons dans la proposition 26.3.

Si  $f$  est une fonction  $\mathbb{C}$ -dérivable représentée par la fonction  $F = P + iQ$ , les équations de Cauchy-Riemann signifient que  $\Delta P = \Delta Q = 0$ , c'est-à-dire que les composantes de la fonction  $f$  sont harmoniques<sup>1</sup>.

PropkwIQwg

**Proposition 26.2** ([1]).

Soit un ouvert  $\Omega \subset \mathbb{C}$ . Nous considérons une application  $f: \mathbb{C} \rightarrow \mathbb{C}$  s'écrivant sous la forme  $f(x + iy) = u(x, y) + iv(x, y)$  où  $u, v: \mathbb{R}^2 \rightarrow \mathbb{R}$  sont des applications réelles.

Nous avons équivalence entre :

---

1. Une fonction  $u$  est harmonique si  $\Delta u = 0$ .

(1)  $f$  est  $\mathbb{C}$ -dérivable.

(2)  $f$  est différentiable et sa différentielle est une similitude<sup>2</sup> (définition 12.333).

(3)  $f$  est différentiable et vérifie

$$\begin{cases} \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \end{cases} \quad (26.6a)$$

$$\begin{cases} \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \end{cases} \quad (26.6b)$$

(4)  $f$  est différentiable et  $\partial_{\bar{z}}f = 0$ .

Si ces conditions sont remplies, nous avons

$$df_{z_0}(z) = (\partial_x u + i\partial_x v)(x_0, y_0)z \quad (26.7)$$

Les équations (26.6) sont les équations de **Cauchy-Riemann**.

*Démonstration.* En plusieurs étapes.

(1) **(1)  $\Leftrightarrow$  (2)** C'est la proposition 12.337.

(2) **(2)  $\Rightarrow$  (3)** Par hypothèse  $f$  est différentiable. Nous allons calculer  $df_{z_0}(z)$ . Nous notons  $z_0 = x_0 + iy_0$  et nous utilisons une des formules du lemme 12.280 :

$$df_{z_0}(x + iy) = \frac{d}{dt} [u(x_0 + tx, y_0 + ty) + iv(x_0 + tx, y_0 + ty)]_{t=0} \quad (26.8a)$$

$$= x(\partial_x u)(x_0, y_0) + y(\partial_y u)(x_0, y_0) + ix(\partial_x v)(x_0, y_0) + iy(\partial_y v)(x_0, y_0) \quad (26.8b)$$

$$= x((\partial_x u)(x_0, y_0) + i(\partial_x v)(x_0, y_0)) + iy((\partial_y v)(x_0, y_0) - i(\partial_y u)(x_0, y_0)) \quad (26.8c)$$

Étant donné que la différentielle est linéaire, les coefficients de  $x$  et de  $iy$  doivent être égaux :

$$(\partial_x u)(x_0, y_0) + i(\partial_x v)(x_0, y_0) = (\partial_y v)(x_0, y_0) - i(\partial_y u)(x_0, y_0). \quad (26.9)$$

En égalisant les parties réelles et imaginaires, nous trouvons les relations (26.6).

Maintenant que nous savons que les deux coefficients dans (26.8c) sont identiques nous avons

$$df_{z_0}(z) = (\partial_x u + i\partial_x v)(x_0, y_0)z, \quad (26.10)$$

c'est-à-dire que nous avons prouvé (26.7) au passage.

(3) **(3)  $\Rightarrow$  (4)** Il suffit de vérifier en utilisant la définition 26.4b, et en remplaçant  $\partial_y u$  et  $\partial_y v$  par ce qu'on peut d'après (26.6) :

$$\partial_{\bar{z}}(u + iv) = \frac{1}{2}(\partial_x + i\partial_y)(u + iv) \quad (26.11a)$$

$$= \frac{1}{2}(\partial_x u + i\partial_x v + i\partial_y u - \partial_y v) \quad (26.11b)$$

$$= \frac{1}{2}(\partial_x u + i\partial_x v - i\partial_x v - \partial_x u) \quad (26.11c)$$

$$= 0. \quad (26.11d)$$

(4) **(4)  $\Rightarrow$  (1)** Nous démontrons (4)  $\Rightarrow$  (2). Nous avons déjà montré que (2) implique (1). Nous posons  $f(x + iy) = u(x, y) + iv(x, y)$  et nous allons calculer  $df_{z_0}(x + iy)$  en reprenant le calcul déjà fait dans (26.8). Pour alléger les notations, nous notons

$$\begin{aligned} \alpha &= (\partial_x u)(x_0, y_0) & \beta &= (\partial_x v)(x_0, y_0) \\ \gamma &= (\partial_y u)(x_0, y_0) & \delta &= (\partial_y v)(x_0, y_0). \end{aligned} \quad (26.12)$$

2. En réalité ce qui est une similitude, c'est la matrice de  $\varphi \circ df_{z_0} \circ \varphi^{-1} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ .

Nous avons :

$$df_{z_0}(x + iy) = x(\alpha + i\beta) + iy(\delta - i\gamma) = \alpha x + \gamma y + i(\beta x + \delta y). \quad (26.13)$$

Donc

$$(\varphi \circ df_{z_0} \circ \varphi^{-1})(x, y) = \begin{pmatrix} \alpha x + \gamma y \\ \beta x + \delta y \end{pmatrix} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (26.14)$$

Par ailleurs, l'hypothèse  $\partial_{\bar{z}} = 0$  donne les équations (26.6), c'est-à-dire  $\delta = \alpha$  et  $\gamma = -\beta$ . La matrice de  $(\varphi \circ df_{z_0} \circ \varphi^{-1})$  est donc  $\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$  qui est une similitude

□

PROPooKXEQooBPjWAH

**Proposition 26.3** ([1]).

Si  $f$  est  $\mathbb{C}$ -dérivable en  $z_0$  nous avons

$$f'(z_0) = (\partial_z f)(z_0). \quad (26.15)$$

*Démonstration.* La proposition 12.337 dit déjà que  $f'(z_0)z = (df_{z_0})(z)$ . Nous devons montrer que  $(\partial_z f)(z_0)z = (df_{z_0})(z)$ . Pour cela nous calculons un peu en utilisant la définition de  $\partial_z$  et les relations (26.6) :

$$(\partial_z f)(z_0) = \frac{1}{2}(\partial_x - i\partial_y)(u + iv)(x_0, y_0) \quad (26.16a)$$

$$= \frac{1}{2}(\partial_x u + i\partial_x v - i\partial_y u + \partial_y v)(x_0, y_0) \quad (26.16b)$$

$$= \frac{1}{2}(\partial_x u + i\partial_x v + i\partial_x v + \partial_x u)(x_0, y_0) \quad (26.16c)$$

$$= (\partial_x u + i\partial_x v)(x_0, y_0). \quad (26.16d)$$

L'équation (26.7) nous permet de terminer :

$$(\partial_z f)(z_0) = (\partial_x u + i\partial_x v)(x_0, y_0) = df_{z_0} \quad (26.17)$$

au sens où  $df_{z_0}$  est l'opérateur sur  $\mathbb{C}$  qui consiste à multiplier par ce qui est à gauche. □

PROPooAGGMooIVQFQB

**Proposition 26.4** (Cauchy-Riemann en coordonnées polaires[1, 695]).

Soit une fonction  $f: \mathbb{C} \rightarrow \mathbb{C}$  que nous supposons être  $\mathbb{C}$ -dérivable dans un voisinage de  $r_0 e^{i\theta_0}$  ( $r_0$  et  $\theta_0$  sont des réels). Nous posons

$$\begin{aligned} \tilde{f}: \mathbb{R} \times \mathbb{R} &\rightarrow \mathbb{C} \\ r, \theta &\mapsto f(re^{i\theta}). \end{aligned} \quad (26.18)$$

Alors

ITEMooRTYYooSTgTAQ

(1) La fonction  $\tilde{f}$  admet des dérivées partielles dans les deux directions en  $(r_0, \theta_0)$ .

ITEMooDHXTooBjxwjY

(2) Les dérivées partielles de  $\tilde{f}$  sont liées par la relation

$$\frac{\partial \tilde{f}}{\partial \theta}(r_0, \theta_0) = ir_0 \frac{\partial \tilde{f}}{\partial r}(r_0, \theta_0). \quad (26.19)$$

ITEMooUUXTooZoDMHI

(3) La fonction  $\tilde{f}$  est de classe  $C^\infty$ .

*Démonstration.* Nous considérons les fonctions réelles  $u$  et  $v$  donnant les parties réelles et imaginaires de  $f$  :

$$f(x + iy) = u(x, y) + iv(x, y). \quad (26.20)$$



Nous avons :

$$\begin{aligned}\frac{\partial \tilde{f}}{\partial r}(r_0, \theta_0) &= \frac{\partial u}{\partial x}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) \cos(\theta_0) \\ &\quad + \frac{\partial u}{\partial y}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) \sin(\theta_0) \\ &\quad + i \frac{\partial v}{\partial x}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) \cos(\theta_0) \\ &\quad + i \frac{\partial v}{\partial y}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) \sin(\theta_0)\end{aligned}\tag{26.21}$$

et

$$\begin{aligned}\frac{\partial \tilde{f}}{\partial \theta}(r_0, \theta_0) &= -\frac{\partial u}{\partial x}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) r_0 \sin(\theta_0) \\ &\quad + \frac{\partial u}{\partial y}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) r_0 \cos(\theta_0) \\ &\quad - \frac{\partial v}{\partial x}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) r_0 \sin(\theta_0) \\ &\quad + i \frac{\partial v}{\partial y}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) r_0 \cos(\theta_0).\end{aligned}\tag{26.22}$$

Nous exprimons tout en termes de  $\partial_y v$  et  $\partial_x v$  en utilisant les équations de Cauchy-Riemann de la proposition 26.2 ainsi que la formule  $\cos(\theta) + i \sin(\theta) = e^{i\theta}$  du lemme 18.11. Pour simplifier les notations, nous notons  $a_0 = (r_0 \cos(\theta_0), r_0 \sin(\theta_0)) \in \mathbb{R}^2$ . Après quelques calculs :

$$(\partial_r \tilde{f})(r_0, \theta_0) = e^{i\theta_0} \frac{\partial v}{\partial y}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) + i e^{i\theta_0} \frac{\partial v}{\partial x}(r_0 \cos(\theta_0), r_0 \sin(\theta_0))\tag{26.23a}$$

$$= e^{i\theta_0} \left( \frac{\partial v}{\partial y}(a_0) + i \frac{\partial v}{\partial x}(a_0) \right),\tag{26.23b}$$

et

$$(\partial_\theta \tilde{f})(r_0, \theta_0) = r_0 e^{i\theta_0} \left( i \frac{\partial v}{\partial y}(a_0) - \frac{\partial v}{\partial x}(a_0) \right)\tag{26.24}$$

En comparant les deux, nous trouvons

$$(\partial_\theta \tilde{f})(r_0, \theta_0) = r_0 i (\partial_r \tilde{f})(r_0, \theta_0).\tag{26.25}$$

□

PROPOCHUEooYsGcQK

### Proposition 26.5.

Si  $f: \mathbb{C} \rightarrow \mathbb{C}$  est holomorphe, alors nous avons

$$df_{z_0} = (\partial_z f)(z_0)\tag{26.26}$$

au sens où l'opérateur linéaire  $df_{z_0}: \mathbb{C} \rightarrow \mathbb{C}$  est l'opération de multiplication par le nombre complexe  $(\partial_z f)(z_0)$ .

*Démonstration.* Soit  $f(x+iy) = f_1(x, y) + i f_2(x, y)$  une fonction holomorphe<sup>3</sup>. Les fonctions réelles  $f_1$  et  $f_2$  sont assujetties aux équations de Cauchy-Riemann de la proposition 26.2 :

$$\begin{cases} \partial_x f_1 = \partial_y f_2 \\ \partial_x f_2 = -\partial_y f_1. \end{cases}\tag{26.27a}$$

$$\tag{26.27b}$$

---

3. Définition 12.331.

Nous avons, en recourant à un petit abus de notation entre  $f_i: \mathbb{R}^2 \rightarrow \mathbb{R}$  et  $f_i: \mathbb{C} \rightarrow \mathbb{R}$  :

$$df_{z_0}(u) = \frac{d}{dt} \left[ f(z_0 + tu) \right]_{t=0} \quad (26.28a)$$

$$= \frac{d}{dt} \left[ f_1(z_0 + tu) + i f_2(z_0 + tu) \right]_{t=0} \quad (26.28b)$$

$$= \partial_x f_1 u_1 + \partial_y f_1 u_2 + i(\partial_x f_2 u_1 + \partial_y f_2 u_2) \quad (26.28c)$$

$$= \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad (26.28d)$$

$$= \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ -\partial_y f_1 & \partial_x f_1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \quad (26.28e)$$

En utilisant le lemme 12.334 nous reconnaissons la matrice de multiplication par le nombre  $\partial_x f_1 - i\partial_y f_1$ . Or justement,

$$\partial_z f = \frac{1}{2} \left( \frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) f = \frac{1}{2} (\partial_x f_1 + i\partial_x f_2 - i\partial_y f_1 + \partial_y f_2), \quad (26.29)$$

qui se réduit à  $\partial_x f_1 - i\partial_y f_1$  lorsque nous y appliquons les équations de Cauchy-Riemann.  $\square$

### 26.1.2 Intégrale sur un chemin dans $\mathbb{C}$

DEFooBPLJooZwsmxi

#### Définition 26.6.

Si nous avons une application  $\gamma: [a, b] \rightarrow \mathbb{C}$  et une fonction  $f: \mathbb{C} \rightarrow \mathbb{C}$ , nous définissons

$$\int_{\gamma} f(z) dz = \int_a^b (f \circ \gamma)(t) \gamma'(t) dt \quad (26.30)$$

pour tous les couples  $(f, \gamma)$  pour lesquels le membre de droite a un sens.

#### 26.7.

Vous noterez que cette définition n'est pas exactement la même que celle 20.40 d'une intégrale curviligne en analyse réelle. Cette dernière demande de prendre la norme de  $\gamma'$ , alors qu'ici nous la gardons telle quelle. D'ailleurs, gardez en tête que  $\gamma$  est une fonction à valeurs dans  $\mathbb{C}$ . Donc  $\gamma'(t)$  peut très bien être un nombre complexe.

PROPOoCUBTooZDcdHX

#### Proposition-Définition 26.8 (Intégrale sur un chemin $C^1$ par morceaux).

Soit un chemin  $C^1$  par morceaux  $\gamma: [a, b] \rightarrow \mathbb{C}$ . Nous considérons deux subdivisions  $a = t_1 < t_2 < \dots < t_n = b$  et  $a = u_1 < u_2 < \dots < u_m = b$  telles que  $\gamma$  soit de classe  $C^1$  sur les intervalles  $]t_i, t_{i+1}[$  et  $]u_j, u_{j+1}[$ . Alors

$$\sum_{i=1}^{n-1} \int_{\gamma: ]t_i, t_{i+1}[ \rightarrow \mathbb{C}} f = \sum_{j=1}^{m-1} \int_{\gamma: ]u_j, u_{j+1}[ \rightarrow \mathbb{C}} f. \quad (26.31)$$

Cette valeur est notée

$$\int_{\gamma} f. \quad (26.32)$$

### 26.1.3 Intégrales sur des chemins fermés

LemtpEOmi

#### Lemme 26.9.

Si  $g$  est une fonction continue dans un ouvert  $\Omega \subset \mathbb{C}$  et si  $g$  admet une primitive complexe sur  $\Omega$  alors

$$\int_{\gamma} g(z) dz = 0 \quad (26.33)$$

pour tout chemin fermé  $\gamma$  de classe  $C^1$  contenu dans  $\Omega$ .

*Démonstration.* Nommons  $G$  une primitive de  $g$ . Par définition,

$$\int_{\gamma} g = \int_{\gamma} G' \quad (26.34a)$$

$$= \int_0^1 G'(\gamma(t))\gamma'(t)dt \quad (26.34b)$$

$$= \int_0^1 (G \circ \gamma)'(t)dt \quad (26.34c)$$

$$= G(\gamma(1)) - G(\gamma(0)) \quad (26.34d)$$

$$= 0 \quad (26.34e)$$

parce que le chemin est fermé :  $\gamma(0) = \gamma(1)$ . □

LemwbwbUR

**Lemme 26.10** (Goursat[597]).

Soit  $\Omega$  un ouvert dans  $\mathbb{C}$  et  $f$  une fonction continue sur  $\Omega$ , holomorphe sur  $\Omega$  moins éventuellement un point (nommé  $z_1 \in \Omega$ ). Soit  $T$ , un triangle<sup>4</sup> fermé inclus dans  $\Omega$ . Alors nous avons

$$\int_{\partial T} f(z)dz = 0. \quad (26.35)$$

*Démonstration.* Nous notons  $\gamma = \partial T$ . Dans la suite nous allons définir une suite de triangles  $T^{(n)}$  et nous noterons  $\gamma_n = \partial T^{(n)}$  avec une orientation que nous allons expliquer. Pour commencer nous posons  $T^{(0)} = T$  et  $\gamma_0 = \partial T^{(0)}$ .

Nous considérons le cas  $z_1 \notin T$ , et nous posons

$$c = l(\gamma)^{-2} \left| \int_{\gamma} f \right|. \quad (26.36)$$

Notre objectif est de montrer que  $c = 0$ . Soit  $A, B, C$  les trois sommets du triangle ; nous divisons le triangle de la façon suivante. D'abord nous considérons les points  $A', B', C'$  respectivement milieux de  $BC$ ,  $AC$  et  $AB$ . En traçant le triangle  $A'B'C'$ , nous construisons quatre triangles que nous nommons  $T_i^{(0)}$ . Le théorème de Thalès<sup>5</sup> assure que le périmètre de chacun des quatre triangles est la moitié du périmètre du grand triangle  $T$ .

Sur  $T$  nous choisissons l'orientation  $ABC$ . De façon à être « compatible », nous choisissons les orientations  $AC'B'$ ,  $BA'C'$  et  $A'CB'$ . La somme de ces trois triangles donne  $T$  plus le triangle  $A'C'B'$ . Par conséquent nous choisissons sur le triangle central l'orientation (inverse)  $A'B'C'$  de façon à avoir

$$\int_{\gamma} f = \sum_{i=1}^4 \int_{\partial T_i^{(0)}} f. \quad (26.37)$$

Cela implique que pour au moins un des quatre triangles (disons  $T_k^{(0)}$  pour fixer les idées) nous ayons

$$\int_{\partial T_k^{(0)}} f \geq \frac{1}{4} \int_{\partial T^{(0)}} f \quad (26.38)$$

Nous notons  $T^{(1)}$  ce triangle. Comme noté précédemment nous avons

$$l(\partial T^{(1)}) = \frac{1}{2} l(\partial T^{(0)}), \quad (26.39)$$

et donc

$$l(\gamma_1)^{-2} \left| \int_{\gamma_1} f \right| = 4l(\gamma_0)^{-2} \left| \int_{\gamma_1} f \right| \geq 4l(\gamma_0)^{-2} \frac{1}{4} \left| \int_{\gamma_0} f \right| = c. \quad (26.40)$$

4. Nous considérons ici le triangle « plein ».

5. Théorème 12.169.

En répétant le procédé nous construisons une suite de triangles  $T^{(n)}$  qui satisfont toujours

$$l(\partial T^{(n)}) = \frac{1}{2^n} l(\partial T^{(0)}). \quad (26.41)$$

Ces triangles forment une suite de fermés emboîtés dont le diamètre tend vers zéro. Leur intersection contient donc exactement un point (lemme 17.125) que nous nommons  $z_0$  (et qui appartient évidemment à  $\Omega$ ). Étant donné que  $f$  est holomorphe nous utilisons le développement limité (26.1) autour de  $z_0$  :

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + s(|z - z_0|)(z - z_0) \quad (26.42)$$

avec  $\lim_{t \rightarrow 0} s(t) = 0$ . Nous posons  $g(z) = f(z_0) + f'(z_0)(z - z_0)$  et nous considérons  $\epsilon > 0$ . Soit  $\alpha > 0$  tel que

$$|f(z) - g(z)| < \epsilon |z - z_0| \quad (26.43)$$

pour tout  $|z - z_0| < \alpha$ . Le  $\alpha$  à choisir pour obtenir cet effet est celui qui donne  $s(|z - z_0|) < \epsilon$ . Soit  $N \in \mathbb{N}$  tel que  $l(\gamma_n) < \alpha$  pour tout  $n > N$ . D'autre part, deux points dans un triangle sont toujours à distance moindre que la longueur d'un côté, donc pour tout  $z \in T^{(n)}$  nous avons  $|z - z_0| < \alpha$  et par conséquent pour tout  $z$  dans  $T^{(n)}$  nous avons

$$|f(z) - g(z)| < \epsilon |z - z_0|. \quad (26.44)$$

Notons que la fonction  $g$  est une dérivée : c'est la dérivée de la fonction

$$G(z) = zf(z_0) + \frac{1}{2}f'(z_0)(z - z_0)^2. \quad (26.45)$$

Par conséquent nous avons

$$\int_{\gamma_n} g = 0 \quad (26.46)$$

par le lemme 26.9. Nous avons donc

$$\left| \int_{\gamma_n} f \right| = \left| \int_{\gamma_n} (f - g) \right| \quad (26.47a)$$

$$\leq l(\gamma_n) \max\{|f(z) - g(z)| \text{ tel que } z \in T^{(n)}\} \quad (26.47b)$$

$$\leq \epsilon l(\gamma_n)^2, \quad (26.47c)$$

et par conséquent

$$c \leq l(\gamma_n)^{-2} \left| \int_{\gamma_n} f \right| \leq \epsilon, \quad (26.48)$$

ce qui signifie que  $c = 0$  parce que  $\epsilon$  est arbitraire. Nous avons donc prouvé le lemme de Goursat dans le cas où le point de non holomorphie  $z_1$  est en dehors de  $T$ .

Si  $z_1$  est sur un côté, disons sur le côté  $AB$ , alors nous considérons un vecteur  $v \in \mathbb{C}$  tel que  $T_\epsilon = T + \epsilon v$  ne contienne  $z_1$  pour aucun  $\epsilon$ . Le vecteur  $v = z_1 - C$  fait par exemple l'affaire. En vertu du point précédent nous avons

$$\int_{\partial T_\epsilon} f = 0 \quad (26.49)$$

pour tout  $\epsilon > 0$ . Étant donné que la fonction  $f$  est continue (y compris en  $z_1$ ), l'intégrale sur  $\partial T$  est également nulle.

Si maintenant le point  $z_1$  est à l'intérieur de  $T$  nous décomposons  $T$  en trois triangles ayant  $z_1$  comme sommet commun. Si nous considérons les orientations  $Az_1C$ ,  $ABz_1$  et  $BCz_1$ , alors nous avons

$$\int_T f = \int_{Az_1C} f + \int_{ABz_1} f + \int_{BCz_1} f, \quad (26.50)$$

alors que par le point précédent les trois intégrales du membre de droite sont nulles.  $\square$

PrpopwQSbJg

**Proposition 26.11** ([597]).

Soient  $\Omega$  un ouvert étoilé et  $f$  une fonction holomorphe sur  $\Omega$  sauf éventuellement en un point  $z_1$  où  $f$  est seulement continue. Alors si  $\gamma$  est un chemin fermé dans  $\Omega$ , nous avons

$$\int_{\gamma} f = 0. \quad (26.51)$$

PropRZCKeO

**Proposition 26.12.**

Si  $f(z) = \sum_n a_n z^n$  a pour rayon de convergence  $R$ , alors  $f$  est  $\mathbb{C}$ -dérivable et nous pouvons dériver terme à terme dans la boule ouverte  $B(0, R)$ .

*Démonstration.* Cela est exactement la proposition <sup>6</sup>15.50. □

### 26.1.4 Homotopie entre applications

DEFooPJKLooCvgxsu

**Définition 26.13** (homotopie entre applications).

Soient des espaces topologiques  $X$  et  $Y$ . Deux applications continues  $f_1, f_2: X \rightarrow Y$  sont **homotopes** si il existe une application continue  $H: [0, 1] \times X \rightarrow Y$  telle que pour tout  $x \in X$  nous avons

$$H(0, x) = f_1(x) \quad (26.52a)$$

$$H(1, x) = f_2(x). \quad (26.52b)$$

LEMooMGFZooGOaGYI

**Lemme 26.14.**

La relation « être homotope à <sup>7</sup> » est une relation d'équivalence sur  $C(X, Y)$ .

*Démonstration.* Pour obtenir  $f \sim f$ , il suffit de prendre  $H(s, x) = f(x)$ .

Si  $f \sim g$ , nous considérons l'homotopie  $H: [0, 1] \times X \rightarrow Y$ . Alors l'application

$$\begin{aligned} M: [0, 1] \times X &\rightarrow Y \\ (t, x) &\mapsto H(1 - t, x) \end{aligned} \quad (26.53)$$

est une homotopie pour  $g \sim f$ .

Si  $f \sim g$  et  $g \sim h$ , nous avons les applications continues  $H$  et  $M$  telles que

$$H(0, x) = f(x) \quad H(1, x) = g(x) \quad (26.54a)$$

$$M(0, x) = g(x) \quad M(1, x) = h(x). \quad (26.54b)$$

L'application

$$\begin{aligned} S: [0, 1] \times X &\rightarrow Y \\ (t, x) &\mapsto \begin{cases} H(2t, x) & \text{si } t \in [0, \frac{1}{2}[ \\ M(2t - 1, x) & \text{si } t \in [\frac{1}{2}, 1]. \end{cases} \end{aligned} \quad (26.55)$$

Nous vérifions que  $S$  est continue en vérifiant la valeur en  $t = 1/2$ . De plus

$$S(0, x) = H(0, x) = f(x) \quad (26.56)$$

et

$$S(1, x) = M(2 - 1, x) = M(1, x) = h(x). \quad (26.57)$$

□

LEMooMJKEooCaVhjD

**Lemme 26.15** ([1]).

Soient des espaces topologiques  $X$  et  $Y$ . Si  $Y$  est connexe par arcs, alors toutes les applications constantes  $X \rightarrow Y$  sont homotopes.

6. Je ne suis pas convaincu que ça s'applique à  $\mathbb{C}$ .

7. Définition 26.13.

*Démonstration.* Soient les applications constantes  $u(x) = u_0$  et  $v(x) = v_0$ . Étant donné que  $Y$  est connexe par arcs, il existe une application continue  $\gamma: [0, 1] \rightarrow Y$  telle que  $\gamma(0) = u_0$  et  $\gamma(1) = v_0$ . Alors l'application

$$\begin{aligned} H: [0, 1] \times X &\rightarrow Y \\ (t, x) &\mapsto \gamma(t) \end{aligned} \quad (26.58)$$

est une homotopie entre les applications  $u$  et  $v$ . □

PROPOoNABDooFtKukO

**Proposition 26.16** ([234]).

Soit un compact  $K$  de  $\mathbb{R}^n$ . Deux applications  $f, g: K \rightarrow \mathbb{C}^*$  sont homotopes dans  $\mathbb{C}^*$  si et seulement si  $f/g$  est homotope à la fonction

$$\begin{aligned} u: K &\rightarrow \mathbb{C} \\ x &\mapsto 1. \end{aligned} \quad (26.59)$$

*Démonstration.* Supposons que  $H: [0, 1] \times K \rightarrow \mathbb{C}^*$  est une homotopie entre  $f$  et  $g$ . Dans ce cas, l'application

$$\begin{aligned} M: [0, 1] \times K &\rightarrow \mathbb{C}^* \\ (t, x) &\mapsto \frac{H(t, x)}{g(x)} \end{aligned} \quad (26.60)$$

est une homotopie entre  $f/g$  et  $u$ . En effet

$$M(0, x) = \frac{H(0, x)}{g(x)} = \frac{f(x)}{g(x)} \quad (26.61)$$

et

$$M(1, x) = \frac{H(1, x)}{g(x)} = 1 = u(x). \quad (26.62)$$

Dans l'autre sens, si  $H$  est une homotopie entre  $f/g$  et  $u$ , alors l'application  $M(t, x) = H(t, x)g(x)$  est une homotopie entre  $f$  et  $g$ . □

### 26.1.5 Intégrale et homotopie

LEMooOFCEooIsuchR

**Lemme 26.17** ([234]).

Soit un rectangle fermé  $R \subset \mathbb{R}^2$ . Si l'application  $H: R \rightarrow \Omega$  est continue, alors pour toute fonction continue  $f: \Omega \rightarrow \mathbb{C}$  nous avons

$$\int_{\partial H} f = 0 \quad (26.63)$$

où  $\partial H$  désigne la frontière de  $H(R)$  dans  $\mathbb{C}$ , et l'intégrale est celle de la définition 26.8.

*Démonstration.* La frontière  $\partial H$  peut être décomposée en 4 parties de classe  $C^1$  :

$$\begin{aligned} (1) \quad & \alpha_1: [0, 1] \rightarrow \Omega \\ & t \mapsto H(t, a), \end{aligned} \quad (26.64)$$

$$\begin{aligned} (2) \quad & \alpha_2: [a, b] \rightarrow \Omega \\ & t \mapsto H(1, t) \end{aligned} \quad (26.65)$$

$$\begin{aligned} (3) \quad & \alpha_3: [0, 1] \rightarrow \Omega \\ & t \mapsto H(1 - t, b) \end{aligned} \quad (26.66)$$

$$\begin{aligned} (4) \quad & \alpha_4: [a, b] \rightarrow \Omega \\ & t \mapsto H(0, b + a - t). \end{aligned} \quad (26.67)$$

Étudions les diverses parties  $\alpha_i$ .

- (1) **Pour  $\alpha_2$**  Première observation :  $\alpha_2 = \gamma_1$ .  
 (2) **Pour  $\alpha_4$**  Nous avons  $\alpha_4 = \gamma_0(b + a - t)$  et donc

$$\int_{\alpha_4} f = \int_a^b (f \circ \alpha_4)(t) \alpha_4'(t) dt = - \int_a^b (f \circ \gamma_0)(b + a - t) \gamma_0'(b + a - t) dt. \quad (26.68)$$

Nous utilisons le changement de variables<sup>8</sup>

$$\begin{aligned} \phi: [a, b] &\rightarrow [a, b] \\ t &\mapsto b + a - t. \end{aligned} \quad (26.69)$$

Le jacobien  $|J_\phi|$  est égal à 1 et donc nous continuons (26.68) de la façon suivante :

$$\int_{\alpha_4} f = - \int_a^b (f \circ \gamma_0)(\phi(t)) \gamma_0'(\phi(t)) dt \quad (26.70a)$$

$$= - \int_a^b (f \circ \gamma_0)(t) \gamma_0'(t) dt \quad (26.70b)$$

$$= - \int_{\gamma_0} f. \quad (26.70c)$$

- (3) **Lien entre  $\alpha_1$  et  $\alpha_3$**  Il y a deux possibilités : soit  $\gamma_1$  et  $\gamma_2$  sont des lacets, soit ce sont des chemins normaux.

- (3a) **Si ils sont des lacets** Alors pour tout  $s$ , l'application

$$\begin{aligned} \Gamma_s: [a, b] &\rightarrow C \\ t &\mapsto H(s, t) \end{aligned} \quad (26.71)$$

est un lacet. En particulier  $H(s, a) = H(s, b)$ . Donc  $\alpha_3(t) = H(1 - t, b) = \alpha_1(1 - t)$ . Dans ce cas nous avons

$$\int_{\alpha_3} f = \int_a^b f(\alpha_3(t)) \alpha_3'(t) dt \quad (26.72a)$$

$$= \int_a^b f(\alpha_1(1 - t)) (-) \alpha_1'(1 - t) dt \quad (26.72b)$$

$$= - \int_a^b (f \circ \alpha_1)(t) dt \quad \text{chm. var. } \phi(t) = 1 - t \quad (26.72c)$$

$$= - \int_{\alpha_1} f. \quad (26.72d)$$

Nous avons alors le calcul

$$0 = \int_{\partial H} f = \int_{\alpha_1} f + \int_{\gamma_1} f + \underbrace{\int_{\alpha_3} f}_{-\int_{\alpha_1} f} - \int_{\gamma_0} f = \int_{\gamma_1} f - \int_{\gamma_0} f. \quad (26.73)$$

Le lemme est prouvé.

- (3b) **Si  $\gamma_0$  et  $\gamma_1$  sont des chemins** Dans ce cas nous avons une homotopie à extrémités fixées. Donc  $\alpha_1$  et  $\alpha_3$  sont des chemins constants, et les intégrales dessus sont nulles et

$$0 = \int_{\partial H} f = \underbrace{\int_{\alpha_1} f}_{=0} + \int_{\gamma_1} f + \underbrace{\int_{\alpha_3} f}_{=0} - \int_{\gamma_0} f = \int_{\gamma_1} f - \int_{\gamma_0} f. \quad (26.74)$$

8. Théorème 14.282(3).

Et le lemme est prouvé. □

THOooVTFXooBgvVyD

**Théorème 26.18** (Cauchy, version homotopique[234, 696]).

Soit un ouvert  $\Omega$  dans  $\mathbb{C}$ . Soit une fonction holomorphe  $f: \Omega \rightarrow \mathbb{C}$ .

Supposons que l'une des deux conditions suivantes soit respectée :

(1) Les chemins  $\gamma_0, \gamma_1: [a, b] \rightarrow \Omega$  sont homotopes à extrémités fixées.

ITEMooESDVooFgVarr

(2) Les lacets  $\gamma_0, \gamma_1: [a, b] \rightarrow \Omega$  sont homotopes.

Alors

$$\int_{\gamma_0} f = \int_{\gamma_1} f. \quad (26.75)$$

*Démonstration.* Considérons l'homotopie  $H: [0, 1] \times [a, b] \rightarrow \Omega$ . Étant donné que  $[0, 1] \times [a, b]$  est un rectangle dans  $\mathbb{R}^2$ , le lemme 26.17 nous indique que

$$\int_{\partial H} f = 0. \quad (26.76)$$

□

CorGZXzuZR

**Corolaire 26.19** ([696]).

Soient  $a \in \mathbb{C}$  ainsi que deux chemins  $\gamma_1$  et  $\gamma_2$  homotopes dans  $\mathbb{C} \setminus \{a\}$ . Alors  $\text{Ind}(\gamma_1, a) = \text{Ind}(\gamma_2, a)$ .

### 26.1.6 Théorème de Tietze (espace normal)

LEMooCLVAooTaNGJk

**Lemme 26.20** ([234]).

Soit un espace topologique normal<sup>9</sup>  $X$ . Soient  $M \in \mathbb{R}^+$  et  $A$  fermé dans  $X$ , et une application continue  $f: A \rightarrow [-M, M]$ .

Il existe une application continue  $g: X \rightarrow [-M/3, M/3]$  telle que pour tout  $x \in A$ ,

$$|f(x) - g(x)| < \frac{2M}{3}, \quad (26.77)$$

*Démonstration.* Nous divisons  $A$  en trois parties :

$$A_- = f^{-1}([-M, -M/3]), \quad (26.78a)$$

$$A_+ = f^{-1}([M/3, M/3]), \quad (26.78b)$$

$$A_0 = f^{-1}([-M/3, M/3]). \quad (26.78c)$$

Les parties  $A_+$  et  $A_-$  sont fermées parce que  $f$  est continue. D'autre part,  $X$  est normal, de telle sorte que le théorème d'Urysohn 21.109 s'applique.

Nous considérons donc une application continue  $g_1: X \rightarrow [0, 1]$  telle que  $g_1^{-1}(A_-) = \{0\}$  et  $g_1^{-1}(A_+) = \{1\}$ . Nous posons alors

$$g(x) = \left(\frac{2M}{3}\right) g_1(x) - \frac{M}{3}. \quad (26.79)$$

Vérification des propriétés de  $g$ .

(1)  $g$  est continue Parce que  $g_1$  est continue.

(2)  $g$  prend ses valeurs dans  $[-M/3, M/3]$  Parce que  $g_1$  prend ses valeurs dans  $[0, 1]$  et que  $t \mapsto (2M/3)t - M/3$  est croissante.

---

9. Définition 7.90.



(3)  $|f(x) - g(x)|$  sur  $A$  Si  $x \in A_-$ , alors  $g_1(x) = 0$  et donc  $g(x) = -M/3$ . Donc

$$|f(x) - g(x)| = |f(x) - M/3| \leq 2M/3 \quad (26.80)$$

parce que  $f(x) \in [-M/3, M/3]$ .

Si  $x \in A_+$ , alors  $g_1(x) = 1$  et donc  $g(x) = M/3$ . Même fin de raisonnement qu'en (26.80).

Si  $x \in A_0$ , alors  $f(x) \in [-M/3, M/3]$  et  $g(x) \in [-M/3, M/3]$  et donc encore  $|f(x) - g(x)| \leq 2M/3$ .

□

LEMooSKSNooEdgFcR

**Lemme 26.21** ([234]).

Soient un espace topologique normal  $X$  ainsi qu'un fermé  $A$  dans  $X$ . Nous considérons une fonction continue  $f: A \rightarrow [-M, M]$ .

Il existe une application  $g: X \rightarrow [-M, M]$  continue prolongeant  $f$ .

*Démonstration.* Nous allons commencer par construire une suite d'applications  $g_i: X \rightarrow \mathbb{R}$  telles que

ITEMooGAIVooQYBCZj

(1) Pour tout  $x \in A$  et pour tout  $n \in \mathbb{N}$ ,

$$|f(x) - \sum_{i=1}^n g_i(x)| \leq \left(\frac{2}{3}\right)^n M \quad (26.81)$$

ITEMooOLNAooEJPdbV

(2) Pour tout  $x \in X$  et pour tout  $i$ ,

$$|g_i(x)| \leq \left(\frac{2}{3}\right)^i \frac{M}{2}. \quad (26.82)$$

Nous commençons par construire  $g_1$  à partir du lemme 26.20 appliqué à la fonction  $f$ . Nous avons donc une application  $g_1: X \rightarrow [-M/3, M/3]$  telle que

$$|f(x) - g_1(x)| \leq \frac{2M}{3} \quad (26.83)$$

EQooGXUIooWnATPw

pour tout  $x \in A$ . Vu que  $g_1$  prend ses valeurs dans  $[-M/3, M/3]$ , elle vérifie la condition (2). De plus (26.83) montre que la condition (1) est vérifiée pour  $n = 1$ .

Et c'est parti pour la récurrence. Nous supposons avoir des applications  $g_i$  pour  $i = 1, \dots, k$  qui vérifient la condition (2), et telles que la condition (1) est satisfaite pour  $n = 1, \dots, k$ . Nous allons maintenant construire  $g_{k+1}$ .

Nous posons

$$h_k: A \rightarrow \mathbb{R} \\ x \mapsto f(x) - \sum_{i=1}^k g_i(x). \quad (26.84)$$

Par hypothèse de récurrence, la fonction  $h_k$  ne prend pas n'importe quelles valeurs dans  $\mathbb{R}$ , mais

$$h_k: A \rightarrow \left[-\left(\frac{2}{3}\right)^k M, \left(\frac{2}{3}\right)^k M\right]. \quad (26.85)$$

Nous construisons  $g_{k+1}$  à partir de ce  $h_k$  et du lemme 26.20. Nous avons donc

$$g_{k+1}: X \rightarrow \left[-\frac{1}{3}\left(\frac{2}{3}\right)^k M, \frac{1}{3}\left(\frac{2}{3}\right)^k M\right] \quad (26.86)$$

vérifiant

$$|h_k(x) - g_{k+1}(x)| \leq \left(\frac{2}{3}\right)^{k+1} M. \quad (26.87)$$

La condition (1) est donc maintenant vérifiée jusqu'à  $n = k + 1$ . Nous vérifions la condition (2) pour  $i = k + 1$ . Simple calcul :

$$|g_{k+1}(x)| \leq \frac{1}{3} \left(\frac{2}{3}\right)^k M = \left(\frac{2}{3}\right)^{k+1} \frac{M}{2}. \quad (26.88)$$

Et voilà pour la définition des applications  $g_i$ .

Vu que  $2/3 < 1$ , la série  $\sum_{i=1}^{\infty} \|g_i\|_{\infty}$  converge normalement (définition 10.110). Notons  $g$  la somme. Le lemme 10.137 donne alors la convergence uniforme  $g_i \xrightarrow{\text{unif}} g$ , et le théorème 12.394 nous assure que  $g$  est continue sur  $X$ .

En ce qui concerne la norme de  $g$ , nous avons, en utilisant la formule (10.403) avec  $q = 2/3$ ,

$$\|g\|_{\infty} \leq \sum_{i=1}^{\infty} \left(\frac{2}{3}\right)^i \frac{M}{2} \leq M. \quad (26.89)$$

Donc  $|g(x)| \leq M$  pour tout  $x \in X$ .

Enfin nous vérifions que  $g$  prolonge  $f$ . Soit  $x \in A$ . Prenez la limite  $n \rightarrow \infty$  dans l'inégalité

$$|f(x) - \sum_{i=1}^n g_i(x)| \leq \left(\frac{2}{3}\right)^n M. \quad (26.90)$$

Nous trouvons que  $|f(x) - g(x)| = 0$ . □

THOooXKGWooFUYlux

**Théorème 26.22** (Théorème de Tietze).

Soit une partie fermée  $A$  de l'espace normal  $X$ . Si la fonction  $f: A \rightarrow \mathbb{R}$  est continue, alors elle se prolonge en une fonction continue  $g: X \rightarrow \mathbb{R}$ .

*Démonstration.* Nous supposons dans un premier temps que  $f$  prenne ses valeurs dans  $] -M, M[$ . À fortiori, elle prend ses valeurs dans  $[-M, M]$  et le lemme 26.21 dit qu'il existe un prolongement continu  $g: X \rightarrow [-M, M]$ . Nous allons construire à partir de là un prolongement continu  $h: X \rightarrow ] -M, M[$  de  $f$ .

Nous posons

$$B = g^{-1}(\{-M, M\}). \quad (26.91)$$

Étant donné que  $f(A) \subset ] -M, M[$ , nous avons  $B \cap A = \emptyset$ . De plus  $A$  est fermé par hypothèse et  $B$  est fermé en tant qu'image réciproque du fermé  $\{-M, M\}$  par l'application continue  $g$ . Nous pouvons donc appliquer le théorème d'Urysohn 21.109.

Nous considérons donc une application  $g_1: X \rightarrow [0, 1]$  telle que  $g_1 = 0$  sur  $B$  et  $g_1 = 1$  sur  $A$ . Enfin nous posons  $h = gg_1$ . Cette application prend ses valeurs dans  $] -M, M[$ , est continue et si  $x \in A$  nous avons

$$h(x) = g(x)g_1(x) = f(x) \times 1 = f(x). \quad (26.92)$$

Ceci règle la question si  $f$  prend ses valeurs dans  $] -M, M[$ .

Nous considérons à présent le cas général  $f: A \rightarrow \mathbb{R}$ . Soit un homéomorphisme  $\phi: \mathbb{R} \rightarrow ] -M, M[$  (par exemple via la proposition 12.203). Nous considérons

$$\begin{aligned} \tilde{f}: A &\rightarrow ] -M, M[ \\ x &\mapsto (\phi \circ f)(x). \end{aligned} \quad (26.93)$$

Nous lui appliquons le premier cas pour avoir une fonction  $\tilde{g}: X \rightarrow ] -M, M[$  qui prolonge  $\tilde{f}$ . Il suffit maintenant de poser

$$\begin{aligned} g: X &\rightarrow \mathbb{R} \\ x &\mapsto (\phi^{-1} \circ \tilde{g})(x). \end{aligned} \quad (26.94)$$

Cela est une application continue et si  $x \in A$ , nous avons

$$g(x) = (\phi^{-1} \circ \tilde{g})(x) = (\phi^{-1} \circ \tilde{f})(x) = (\phi^{-1} \circ \phi \circ f)(x) = f(x). \quad (26.95)$$

□

**Corolaire 26.23.**

Soit un fermé  $A$  dans un espace normal  $X$ . Si  $f: A \rightarrow \mathbb{C}$  est continue, alors elle se prolonge en une fonction continue  $g: X \rightarrow \mathbb{C}$ .

*Démonstration.* Les parties réelles et imaginaires de  $f$  sont continues. Il suffit de leur appliquer le théorème de Tietze 26.22.  $\square$

**26.2 Logarithme complexe****26.2.1 La fonction argument**

Nous savons la définition 18.6 de l'exponentielle complexe.

**Définition 26.24.**

Un **logarithme** de  $\alpha \in \mathbb{C}$  est une solution de l'équation  $e^z = \alpha$ .

Notons bien que cela définit un logarithme, et non le logarithme.

LEMooUMESooJVzeDb

**Lemme 26.25.**

Si  $z_1$  et  $z_2$  sont des logarithmes de  $\alpha$  alors il existe  $k \in \mathbb{Z}$  tel que  $z_1 = z_2 + 2ik\pi$ .

*Démonstration.* Nous commençons par déterminer les logarithmes de  $\alpha = 1$ . Nous avons besoin de  $e^{a+bi} = 1$  ( $a, b \in \mathbb{R}$ ). Nous avons

$$e^a e^{bi} = 1, \quad (26.96)$$

et en prenant la norme nous trouvons  $|e^a| = 1$ , ce qui donne  $a = 0$ . Ensuite  $e^{bi} = 1$ , qui signifie  $b = 2k\pi$ . Les logarithmes de 1 sont donc les nombres de la forme  $2ik\pi$ .

Soient maintenant  $z_1$  et  $z_2$  des logarithmes de  $\alpha$ . Alors  $e^{z_1} = e^{z_2}$ , donc<sup>10</sup>  $e^{z_1 - z_2} = 1$ , ce qui signifie que  $z_1 - z_2$  est un logarithme de 1. Donc il existe un  $k \in \mathbb{Z}$  tel que  $z_1 - z_2 = 2ik\pi$ .  $\square$

**Remarque 26.26.**

Jusqu'ici nous n'avons pas donné de conditions donnant l'existence d'un logarithme. Nous avons seulement supposé des existences et donné des propriétés sur ces hypothétiques objets.

**Définition 26.27 ([697]).**

Si  $z \in \mathbb{C}^*$  nous définissons la **valeur principale** de son argument le nombre  $\theta \in ]-\pi, \pi]$  tel que

$$z = |z|e^{i\theta} \quad (26.97)$$

Nous le notons  $\arg(z)$ .

NORMooOGHNooYriCBH

**26.28.**

Il ne faut pas se ruer sur  $\arg(x + iy) = \arctan(y/x)$ . Pour rappel, la fonction  $\arctan$  a été définie dans le théorème 18.39, et elle prend ses valeurs dans  $]-\pi/2, \pi/2[$ . La formule  $v(x, y) = \arctan(y/x)$  n'est donc valable que pour  $x > 0$ . Les valeurs sont :

$$\arg(x + iy) = \begin{cases} \arctan(y/x) & \text{si } x > 0 \\ \pi + \arctan(y/x) & \text{si } x < 0 \text{ et } y \geq 0 \\ -\pi + \arctan(y/x) & \text{si } x < 0 \text{ et } y < 0 \\ \frac{\pi}{2} & \text{si } x = 0 \text{ et } y > 0 \\ -\frac{\pi}{2} & \text{si } x = 0 \text{ et } y < 0. \end{cases} \quad (26.98)$$

EQooPJVFooSEKToy

Pour  $x > 0$  nous avons  $\arg(x + iy) = \arctan(y/x)$  parce que justement la fonction  $\arctan$  prend ses valeurs en particulier entre  $-\pi$  et  $\pi$ . Pour  $x < 0$  et  $y > 0$  nous avons  $\arg(x + iy) = \pi + \arctan(y/x)$  (dans ce cas,  $\arctan(y/x) < 0$ ) et si  $x < 0$ ,  $y < 0$  nous avons  $\arg(x + iy) = -\pi + \arctan(y/x)$ .

10. C'est facile de dire « donc ». Il faut surtout citer la proposition 18.9(2).

NORMooMRBEooVtTeIA

**26.29** (Les dérivées partielles de la fonction argument).

Vu que nous en aurons besoin plusieurs fois, nous calculons maintenant les dérivées partielles de la fonction

$$\begin{aligned}\varphi: \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \arg(x + iy).\end{aligned}\tag{26.99}$$

Nous commençons par la dérivée  $\partial_x \varphi(x, y)$ . Et il y a de nombreux cas à séparer.

(1)  $x > 0$

Nous avons

$$\frac{\partial \varphi}{\partial x}(x, y) = \lim_{\epsilon \rightarrow 0} \frac{\arctan(y/(x + \epsilon)) - \arctan(y/x)}{\epsilon},\tag{26.100}$$

qui n'est autre que la dérivée de la fonction  $x \mapsto \arctan(y/x)$ . Nous pouvons la calculer facilement avec le théorème 18.39(2) :

$$\frac{\partial \varphi}{\partial x}(x, y) = -\frac{y}{x^2 + y^2}.\tag{26.101}$$

(2)  $x < 0$

Nous avons

$$\frac{\partial \varphi}{\partial x}(x, y) = \lim_{\epsilon \rightarrow 0} \frac{\pm \pi + \arctan(y/(x + \epsilon)) - (\pm \pi + \arctan(y/x))}{\epsilon}\tag{26.102}$$

où les signes  $\pm$  dépendent du signe de  $y$ . De toutes façons, les termes en  $\pi$  se simplifient et le calcul est le même que celui du cas  $x > 0$ . Encore une fois nous avons

$$\frac{\partial \varphi}{\partial x}(x, y) = -\frac{y}{x^2 + y^2}.\tag{26.103}$$

(3)  $x = 0$

Nous devons calculer

$$\frac{\partial \varphi}{\partial x}(0, y) = \lim_{\epsilon \rightarrow 0} \frac{\arg(\epsilon + iy) - \arg(iy)}{\epsilon}.\tag{26.104}$$

Il y a quatre cas d'après les signes de  $\epsilon$  (séparer limite à gauche et à droite) et  $y$ .

Si  $\epsilon > 0$  et  $y > 0$  alors nous avons à faire le calcul

$$\lim_{\epsilon \rightarrow 0^+} \frac{\arctan(y/\epsilon) - \pi/2}{\epsilon}\tag{26.105}$$

qui se traite par la règle de l'Hospital. Cela donne  $-1/y$ .

Les trois autres cas ne se distinguent que par des constantes au numérateur, lesquelles disparaissent en appliquant la règle de l'Hospital 11. Au final,

$$\frac{\partial \varphi}{\partial x}(0, y) = -\frac{1}{y}.\tag{26.106}$$

Nous avons calculé jusqu'ici :

$$\frac{\partial \varphi}{\partial x}(x, y) = \frac{-y}{x^2 + y^2}\tag{26.107}$$

EQooAOJPooQryUBR

pour tout  $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ . En particulier vous avez noté que cette dérivée partielle est continue sur  $\mathbb{R}^2 \setminus \{(0, 0)\}$ .

Nous calculons à présent la dérivée partielle par rapport à  $y$  :

$$\frac{\partial \varphi}{\partial y}(x, y) = \lim_{\epsilon \rightarrow 0} \frac{\arg(x + iy + i\epsilon) - \arg(x + iy)}{\epsilon}.\tag{26.108}$$

11. Nonobstant le fait que ces constantes se mettent bien pour avoir un vrai cas d'indétermination  $0/0$ , sinon la règle de l'Hospital ne s'applique pas.

(1)  $x > 0$ 

Nous avons à calculer

$$\lim_{\epsilon \rightarrow 0} \frac{\arctan \frac{y+\epsilon}{x} - \arctan \frac{y}{x}}{\epsilon}, \quad (26.109)$$

qui n'est autre que la dérivée de la fonction  $t \mapsto \arctan \frac{t}{x}$  en  $t = y$ . Résultat :

$$\frac{\partial \varphi}{\partial y}(x, y) = \frac{x}{x^2 + y^2}. \quad (26.110)$$

(2)  $x < 0$  et  $y \neq 0$ 

Le calcul à faire est :

$$\lim_{\epsilon \rightarrow 0} \frac{\pm\pi + \arctan \frac{y+\epsilon}{x} - (\pm\pi + \arctan \frac{y}{x})}{\epsilon} \quad (26.111)$$

Une chose importante à remarquer est que dans le calcul de la limite nous pouvons supposer que  $y$  et  $y + \epsilon$  aient le même signe, quelle que soit la valeur et le signe de  $\epsilon$  (assez petit). C'est pour cela que les deux termes  $\pm\pi$  arrivent avec le même signe des deux côtés de la différence, et se simplifient. Nous tombons sur une limite déjà faite et

$$\frac{\partial \varphi}{\partial y}(x, y) = \frac{x}{x^2 + y^2} \quad (26.112)$$

(3)  $x < 0$  et  $y = 0$ Vu que  $x < 0$  nous avons  $\arg(x) = \pi$  et nous devons calculer

$$\lim_{\epsilon \rightarrow 0} \frac{\arg(x + i\epsilon) - \pi}{\epsilon}. \quad (26.113)$$

La limite  $\epsilon \rightarrow 0^+$  est classique et donne  $1/x$ .Mais la limite  $\epsilon \rightarrow 0^-$  n'existe pas :

$$\lim_{\epsilon \rightarrow 0^-} \frac{-\pi + \arctan(\epsilon/x) - \pi}{\epsilon} \quad (26.114)$$

n'existe pas.

Donc

$$\frac{\partial \varphi}{\partial y}(x, 0) \quad (26.115)$$

n'existe pas pour  $x < 0$ .(4)  $x = 0$  et  $y \neq 0$ 

Le calcul est immédiat

$$\lim_{\epsilon \rightarrow 0} \frac{\arg(iy + i\epsilon) - \arg(iy)}{\epsilon} = 0, \quad (26.116)$$

donc

$$\frac{\partial \varphi}{\partial y}(0, y) = 0. \quad (26.117)$$

En ce qui concerne la continuité, nous avons que  $\partial_y \varphi$  est continue partout sauf sur la demi-droite  $\{(x, 0) \text{ tel que } x \leq 0\}$  où elle n'existe pas.

### 26.2.2 Une définition possible du logarithme

DEFooWDYNooYIXVMC

#### Définition 26.30.

Nous définissons la fonction **logarithme** par

$$\begin{aligned} \ln: \mathbb{C}^* &\rightarrow \mathbb{C} \\ z &\mapsto \ln(|z|) + i \arg(z) \end{aligned} \quad (26.118)$$

où le  $\ln$  à droite est le logarithme usuel sur  $\mathbb{R}^+$ .

**Remarque 26.31.**

Cette fonction généralise le logarithme déjà vu sur  $]0, \infty[ \subset \mathbb{R}$ . En effet pour des valeurs de  $z$  dans cette partie nous avons  $\arg(z) = 0$  et  $|z| = z$ .

**Lemme 26.32.**

Le nombre  $\ln(z)$  est un logarithme de  $z$ .

*Démonstration.* Nous avons

$$e^{\ln(z)} = e^{\ln|z|} e^{i \arg(z)} = |z| e^{i \arg(z)} = z. \quad (26.119)$$

Nous avons utilisé le fait que  $e^{\ln(x)} = x$  pour  $x \in \mathbb{R}^+$  et  $|z| e^{i \arg(z)} = z$  par définition de la fonction  $\arg$ .  $\square$

Notons que si on avait pris d'autres conventions pour définir  $\arg$ , nous aurions eu d'autres définitions possibles de  $\ln$ .

**Exemple 26.33.**

Nous avons

$$\ln(-1) = \ln(1) + i \arg(-1). \quad (26.120)$$

Mais  $\ln(1) = 0$  et  $\arg(-1) = \pi$  (et non  $-\pi$ ), donc

$$\ln(-1) = i\pi. \quad (26.121)$$

C'est cette définition du logarithme qui est prise par Sage, et c'est cela qui lui permet de donner la primitive de  $1/x$  comme  $\ln(x)$  et non  $\ln(|x|)$ , parce que Sage connaît les logarithmes de nombres réels négatifs :

```

1
2 SageMath version 7.3, Release Date: 2016-08-04
3 Type "notebook()" for the browser-based notebook interface.
4 Type "help()" for help.
5
6 sage: ln(-1)
7 I*pi
8 sage: f(x)=1/x
9 sage: f.integrate(x)
10 x |--> log(x)

```

tex/sage/sageSnip010.sage

$\triangle$

Nous avons jusqu'ici défini une fonction sur  $\mathbb{C}^*$  qui fait correspondre à chaque nombre complexe un de ses logarithmes. Il reste quelques questions à régler :

- Est-ce que cette fonction est continue ? Holomorphe ? (réponses : non et non)
- Si non, est-ce qu'il y avait moyen de trouver une définition plus efficace ? (réponse : non)

LEMooMUOIooCnoWwq

**Lemme 26.34.**

La fonction  $\ln$  n'est pas continue sur  $]-\infty, 0]$ .

*Démonstration.* Attention à bien comprendre l'énoncé. La fonction

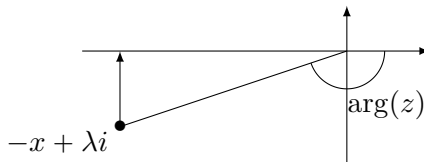
$$\begin{aligned} f: ]-\infty, 0[ &\rightarrow \mathbb{C} \\ x &\mapsto \ln(x) \end{aligned} \quad (26.122)$$

est continue. D'ailleurs c'est  $\ln(x) = \ln(|x|) + i\pi$ . Ce dont il est question dans l'énoncé, c'est de la fonction  $\ln$  vue comme fonction sur  $\mathbb{C}^*$ .

Soit  $x > 0$  dans  $\mathbb{R}$ ; nous avons

$$\ln(-x) = \ln(x) + i\pi. \quad (26.123)$$

Cependant  $\lim_{\substack{\lambda \rightarrow 0^- \\ \lambda \in \mathbb{R}}} \ln(-x + \lambda i)$  va valoir  $\ln(|x|) - i\pi$ . En effet lorsque  $\lambda < 0$  est petit, l'argument de  $-x + \lambda i$  se rapproche de  $-\pi$  (et non de  $\pi$ ).



Donc

$$\lim_{\substack{\lambda \rightarrow 0^- \\ \lambda \in \mathbb{R}}} \ln(-x + \lambda i) = \lim_{\lambda \in \mathbb{R}} \ln(|x + \lambda i|) + i \arg(-x + \lambda i) = \ln(|x|) - i\pi. \quad (26.124)$$

Nous n'avons donc pas continuité de la fonction logarithme comme fonction sur  $\mathbb{C}^*$ .  $\square$

THOooWUXOooYKvLbJ

### Théorème 26.35.

La restriction

$$\ln: \mathbb{C} \setminus ]-\infty, 0] \rightarrow \mathbb{C} \quad (26.125)$$

est holomorphe.

*Démonstration.* Nous allons utiliser la proposition 12.337 et considérer la fonction

$$\begin{aligned} F: S &\rightarrow \mathbb{R}^2 \\ (x, y) &\mapsto (\ln(|x + iy|), \arg(x + iy)) \end{aligned} \quad (26.126)$$

où  $S = \mathbb{R}^2 \setminus \{(x, 0) \text{ tel que } x \leq 0\}$ . Nous devons vérifier que  $F$  est différentiable et que sa différentielle en un point de  $S$  est une similitude.

Nous posons

$$u(x, y) = \ln(\sqrt{x^2 + y^2}) \quad (26.127)$$

et

$$v(x, y) = \arg(x + iy). \quad (26.128)$$

Les dérivées partielles de  $u$  ne sont pas très compliquées :

```

1 sage: var('x, y')
2 (x, y)
3 sage: u(x, y) = ln(sqrt(x**2 + y**2))
4 sage: u.diff(x)
5 (x, y) |--> x/(x^2 + y^2)

```

tex/sage/sageSnip011.sage

c'est-à-dire

$$\frac{\partial u}{\partial x} = \frac{x}{x^2 + y^2} \quad (26.129a)$$

$$\frac{\partial u}{\partial y} = \frac{y}{x^2 + y^2}. \quad (26.129b)$$

Pour celles de  $v$  par contre, il faut se poser des questions, par exemples résister à la tentation d'écrire  $v(x, y) = \arctan(y/x)$  et lire 26.28.

Nous avons déjà calculé les dérivées partielles de  $v$  dans 26.29, et nous avons vu qu'elles étaient continues sur  $\mathbb{R}^2$  privé de la demi-droite.

Vu que les dérivées partielles sont continues, le théorème 12.320 nous dit que  $F$  est différentiable. La matrice de la différentielle est alors la matrice des dérivées partielles

$$\begin{pmatrix} \frac{x}{x^2+y^2} & \frac{y}{x^2+y^2} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}, \quad (26.130)$$

qui a la forme requise (12.910) pour que la proposition 12.337 nous assure que  $\ln$  soit  $\mathbb{C}$ -dérivable, c'est-à-dire holomorphe.  $\square$

### 26.2.3 Pas plus de continuité

Bon. La fonction logarithme que nous avons définie est holomorphe sur  $\mathbb{C}^*$  privé d'une demi-droite

$$U = \{z \in \mathbb{C} \text{ tel que } \operatorname{Im}(z) = 0, \operatorname{Re}(z) \leq 0\}. \quad (26.131)$$

Et elle n'est pas continue sur  $U$ ; elle y est cependant continue « par le haut ». Pouvons-nous faire mieux? Nous allons maintenant prouver quelques résultats d'impossibilité de faire mieux que holomorphe partout sauf une partie pas si petite que ça.

#### Proposition 26.36.

Il n'existe pas de fonctions continues  $f: \mathbb{C}^* \rightarrow \mathbb{C}$  telle que  $e^{f(z)} = z$  pour tout  $z \in \mathbb{C}^*$ .

*Démonstration.* Pour tout  $z$ , le nombre  $f(z)$  est un logarithme de  $z$ . Or  $\ln(z)$  en est également un. Donc par le lemme 26.25

$$f(z) = \ln(z) + 2ik(z)\pi \quad (26.132)$$

pour une certaine fonction  $k: \mathbb{C}^* \rightarrow \mathbb{Z}$ . Sur le domaine d'holomorphie de  $\ln$ , les fonctions  $\ln$  et  $f$  étant continues, la fonction  $k$  l'est aussi. Mais une fonction continue à valeurs dans  $\mathbb{Z}$  est constante (son domaine est connexe).

Il existe donc  $k \in \mathbb{Z}$  tel que

$$f(z) = \ln(z) + 2ik\pi \quad (26.133)$$

au moins pour tout  $z \in \mathbb{C}^* \setminus U$ . Une telle fonction ne peut pas être continue sur  $U$  parce que  $\ln$  ne l'est pas.  $\square$

Ok. Pas continue sur tout  $\mathbb{C}$ . Mais continue sur un peu plus que  $\mathbb{C}$  privé de toute une demi-droite? La proposition suivante répond que bof.

#### Proposition 26.37.

Soit  $\Omega$  un ouvert de  $\mathbb{C}$  contenant  $S(0, r)$  (le cercle centré en 0 et de rayon  $r > 0$ ). Il n'existe pas de fonction continue  $f: \Omega \rightarrow \mathbb{C}$  telle que  $e^{f(z)} = z$  pour tout  $z \in \Omega$ .

*Démonstration.* Encore une fois, pour tout  $z \in \Omega$  nous avons

$$f(z) = \ln(z) + 2i\pi k(z) \quad (26.134)$$

pour une certaine fonction  $k: \Omega \rightarrow \mathbb{Z}$ . Nous considérons la demi-droite  $U$  de (26.131). Sur  $\Omega \setminus U$ , la fonction  $\ln$  est continue et  $k$  doit également l'être. Donc  $k$  est constante sur les composantes connexes de  $\Omega \setminus U$ .

Vu que  $S(0, r)$  est compact, on peut le recouvrir par un nombre fini de boules centrées en des points de  $S(0, r)$ . En prenant le minimum des rayons de ces boules, nous voyons que  $\Omega$  contient une couronne

$$\{z \in \mathbb{C} \text{ tel que } r - \delta \leq |z| \leq r + \delta\}. \quad (26.135)$$

Soit le point  $x_0 = -r$ . C'est un point de  $\Omega$  contenu dans  $U$ . Nous allons prouver que  $B(x_0, \delta) \setminus U$  est dans une seule composante connexe de  $\Omega$ .



Soit un point  $z_1 \in B(x_0, \delta)$  situé au-dessus de  $U$ , et  $z_2$  un point de  $B(x_0, \delta)$  situé en dessous de  $U$ . Le cercle  $S(0, r)$  coupe  $B(x_0, \delta)$  en deux points : un au-dessus et un en-dessous de  $U$ . On peut lier  $z_1$  au point de « sortie » supérieur de  $S(0, r)$  en restant dans  $B(x_0, \delta)$  ; ce point est ensuite relié en suivant le cercle au point d'entrée inférieur du cercle dans  $B(x_0, \delta)$ . Ce dernier point est lié à  $z_2$  par un chemin restant dans la boule.

Tout cela pour dire que  $z_1$  et  $z_2$  sont dans la même composante connexe de  $\Omega$  et que  $k(z_1) = k(z_2)$ . Il existe donc  $k \in \mathbb{Z}$  tel que

$$f(z) = \ln(z) + 2ik\pi \quad (26.136)$$

sur  $B(x_0, \delta) \setminus U$ . Une telle fonction  $f$  ne peut pas être continue.  $\square$

## 26.2.4 Pas d'unicité : autres déterminations de l'argument

NORMooFCDooFDzAjp

### 26.38.

Nous avons pris la fonction d'argument  $\arg: \mathbb{C} \rightarrow ]-\pi, \pi]$ . Il y en a évidemment beaucoup d'autres de possibles. Par exemple pour  $\alpha \in \mathbb{R}$  nous pouvons considérer

$$\arg_{\alpha+}: \mathbb{C} \rightarrow ]\alpha, \alpha + 2\pi] \quad (26.137)$$

EQooNKKDooOyJxXe

ou

$$\arg_{\alpha-}: \mathbb{C} \rightarrow [\alpha, \alpha + 2\pi[. \quad (26.138)$$

En posant

$$\ln_{\alpha\pm}(z) = \ln(|z|) + i \arg_{\alpha\pm}(z) \quad (26.139)$$

nous avons une fonction réciproque de l'exponentielle définie sur  $\mathbb{C}^*$  et holomorphe sur  $\mathbb{C}^*$  privé d'une demi-droite  $D_\alpha$  (dépendante de la valeur de  $\alpha$ ).

La différence entre  $\ln_{\alpha+}$  et  $\ln_{\alpha-}$  est seulement la valeur sur la demi-droite de non-holomorphie. L'une sera semicontinue d'un côté et l'autre, de l'autre côté.

### Remarque 26.39.

La fonction  $\arg_{0-}$  a déjà été utilisée en 18.5.2 pour écrire un inverse de la fonction

$$\begin{aligned} \varphi: [0, 2\pi[ &\rightarrow S^1 \\ t &\mapsto e^{it}. \end{aligned} \quad (26.140)$$

### Définition 26.40 ([650]).

Soit un ouvert  $\Omega \subset \mathbb{C}^*$ . Nous disons que la fonction  $f: \Omega \rightarrow \mathbb{C}$  est une **détermination** sur  $\Omega$  si elle est continue et vérifie

$$e^{f(z)} = z \quad (26.141)$$

pour tout  $z \in \Omega$ .

Les différents résultats vus jusqu'ici montrent qu'il n'existe pas de détermination du logarithme sur  $\mathbb{C}^*$ .

### Définition 26.41.

La **détermination principale** du logarithme est la restriction de notre logarithme 26.30

$$\begin{aligned} \ln: \mathbb{C}^* &\rightarrow \mathbb{C} \\ z &\mapsto \ln(|z|) + i \arg(z) \end{aligned} \quad (26.142)$$

à l'ouvert  $\mathbb{C}^* \setminus U$  où  $U$  est la demi-droite  $\operatorname{Re}(z) \leq 0, \operatorname{Im}(z) = 0$  de  $\mathbb{C}$ .

REMooFBLooDnkmjR

### Remarque 26.42.

Beaucoup de sources[698] ne définissent pas  $\ln_{\alpha\pm}$  sur la droite  $D_\alpha$ . C'est-à-dire qu'ils notent  $\ln_\alpha$  notre fonction  $\ln_{\alpha+}$  restreinte à  $\mathbb{C}^* \setminus D_\alpha$ . Dans ce cas, les fonctions  $\ln_{\alpha+}$  et  $\ln_{\alpha-}$  sont identiques<sup>12</sup>.

12. Cela n'est pas tout à fait évident ; vous devriez y penser.

Cette remarque est importante parce que certains vont vous dire « le logarithme n'est pas défini sur la demi-droite » ; de leur point de vue, la fonction que nous avons définie est une prolongation (non continue) à  $U$  du logarithme, qui est continu.

- (1) Certaines personnes pourraient vous dire que notre logarithme « n'est pas bien défini parce que si on fait le tour dans un sens ou dans l'autre nous n'obtenons pas la même valeur pour  $\ln(z)$  lorsque  $z$  est sur  $U$  ». Et cela avec des arguments aussi forts que «  $2\pi$  et  $0$ , c'est le même point ».

Nous préférons être bien clairs<sup>13</sup> sur ce point : notre fonction  $\ln$  est parfaitement définie sur  $\mathbb{C}^*$  et  $2\pi$  n'est pas la même chose que zéro. En particulier  $\arg(e^{2i\pi}) = 0$  et  $\arg(e^{-i\pi}) = \pi$  et non  $-\pi$ .

- (2) Il n'en reste pas moins que Sage donne  $\ln(-1) = i\pi$  et que nous avons choisi de faire de même, parce que le Frido n'est pas un cours d'agrégation, mais un texte qui donne quelques éléments de mathématique dans le but d'utiliser Sage efficacement.
- (3) Tout ceci pour dire que si vous utilisez ce livre pour l'agrégation, vous devriez sérieusement considérer l'option de ne pas donner du logarithme la définition donnée ici, mais bien sa restriction.

En fait notre logarithme est maximum pour la propriété « être une réciproque de l'exponentielle » alors que beaucoup de monde préfère avoir une fonction maximale pour la propriété « être réciproque de l'exponentielle tout en étant continue ».

De toutes les fonctions ayant le droit de vouloir être appelée « logarithme », celle que nous avons choisie (un peu arbitrairement) pour s'appeler « logarithme » et accaparer de la notation «  $\ln$  » est  $\ln_{\pi+}$ . Elle est d'une certaine manière celle qui arrive le plus naturellement.

En effet si nous pensons au logarithme népérien  $\ln: \mathbb{R}^+ \rightarrow \mathbb{R}$  que nous voulons prolonger sur  $\mathbb{R}$ , nous devons poser

$$\ln(-x) = \ln(-1) + \ln(x) \quad (26.143)$$

pour  $x > 0$ . Que peut valoir  $\ln(-1)$  ? Il doit vérifier  $e^{\ln(-1)} = -1$ . La première valeur qui nous tombe sous la main est  $\ln(-1) = \pi$ . Bien entendu, d'autres possibilités existent, comme  $\ln(-1) = 2017\pi$  par exemple.

### 26.2.5 Pas d'unicité : développement en série

Pour  $z_0 \in \mathbb{C}^*$  nous pouvons écrire un développement en série de la réciproque de l'exponentielle autour de  $z_0$ . La fonction ainsi définie est holomorphe sur la boule  $B(z_0, |z_0|)$  et diverge en dehors de cette boule.

Voilà encore une fonction « logarithme » pour chaque point de  $\mathbb{C}^*$ . Nous nommons  $\ln_{z_0}$  la fonction

$$\ln_{z_0}: B(z_0, |z_0|) \rightarrow \mathbb{C} \quad (26.144)$$

donnée par la série.

En général nous n'avons pas  $\ln_{z_1} = \ln_{z_2}$  sur l'intersection des disques de convergence. Si c'était le cas, de proche en proche nous pourrions construire une fonction continue réciproque du logarithme sur  $\mathbb{C}^*$ , ce qui est impossible.

### 26.2.6 Pas d'unicité : laquelle choisir ?

Bon. Pour chaque demi-droite  $D$  nous avons une détermination du logarithme sur  $\mathbb{C}^* \setminus D$ . Et pour tout  $z_0 \in \mathbb{C}^*$  nous en avons une sur  $B(z_0, |z_0|)$ .

En pratique, quel logarithme choisir ? Cela dépend du problème.

Si vous avez besoin ou envie de travailler avec des série entières, le mieux est de choisir une détermination donnée par un développement autour d'un point bien choisi par rapport à votre problème.

---

13. Est-ce qu'il faut vraiment un pluriel ici ?

Si vous avez surtout besoin d'holomorphicité, et que vous en avez besoin sur un grand domaine, vous devriez choisir une détermination sur un des ensembles  $\mathbb{C}^* \setminus D_\alpha$  en choisissant  $\alpha$  de telle sorte que la demi-droite maudite ne passe pas par la zone sur laquelle vous travaillez.

Dans tous les cas, vous devez préciser très explicitement la détermination choisie. Dans ce texte, sauf mention du contraire, nous utiliserons la détermination principale, et même son extension (non continue) à  $\mathbb{C}^*$ . Lorsque nous aurons besoin d'holomorphicité, nous préciserons que nous considérons la restriction.

### 26.2.7 Logarithme comme primitive

Tout le monde sait <sup>14</sup> que le logarithme  $\ln: \mathbb{R}^+ \rightarrow \mathbb{R}$  est une primitive de la fonction  $x \mapsto 1/x$ . Qu'en est-il dans le cas complexe ? Tout d'abord précisons que nous ne comptons pas encore parler d'intégrale sur  $\mathbb{C}$ , mais seulement d'intégrales sur  $\mathbb{R}$  d'une fonction à valeur complexes.

PROPOOIJVooKueuYJ

#### Proposition 26.43.

Si  $z \in \mathbb{C}$  alors

$$\int \frac{1}{x+z} dx = \ln(x+z) \quad (26.145)$$

EQooAHYXooTPGXDS

*Démonstration.* Il est important de comprendre que la formule (26.145) est un abus de notation pour dire que si nous considérons la fonction

$$\begin{aligned} \varphi: \mathbb{R} &\rightarrow \mathbb{C} \\ x &\mapsto \ln(x+z) \end{aligned} \quad (26.146)$$

alors nous avons  $\varphi'(x) = \frac{1}{x+z}$ . Ici la dérivation est une dérivation sur  $\mathbb{R}$  et l'intégrale est une intégrale sur  $\mathbb{R}$ , c'est-à-dire « composante par composantes ». La fonction  $\varphi$  se décompose en partie réelle et imaginaire qui sont à dériver séparément :

$$\varphi(x) = \ln(|x+z|) + i \arg(x+z). \quad (26.147)$$

#### (1) Si $z$ est imaginaire pur

Nous posons  $z = \lambda i$  avec  $\lambda \in \mathbb{R}^*$ . D'abord nous avons

$$\frac{1}{x+\lambda i} = \frac{x}{x^2+\lambda^2} - i \frac{\lambda}{x^2+\lambda^2}. \quad (26.148)$$

La partie réelle de  $\varphi(x)$  est

$$\varphi_1(x) = \ln(\sqrt{x^2+\lambda^2}), \quad (26.149)$$

dont la dérivée est

$$\varphi_1'(x) = \frac{x}{x^2+\lambda^2}, \quad (26.150)$$

qui correspond bien à la partie réelle de  $\frac{1}{x+\lambda i}$ .

En ce qui concerne la partie imaginaire,  $\varphi_2(x) = \arg(x+\lambda i)$ , et sa dérivée n'est rien d'autre que la dérivée partielle par rapport à  $x$  de la fonction argument, déjà calculée en (26.107) :

$$\varphi_2'(x) = \frac{-\lambda}{x^2+\lambda^2}. \quad (26.151)$$

Cela est bien la partie imaginaire de  $\frac{1}{x+\lambda i}$ .

Notons que nous n'avons pas de problèmes sur la demi-droite des réels négatifs parce que nous ne considérons au final que la dérivée partielle par rapport à  $x$  de la fonction argument, laquelle existe et est continue, même sur cette partie.

14. Proposition 15.93.

(2) Pour  $z$  quelconque

Soit  $z = s + \lambda i$  avec  $s, \lambda \in \mathbb{R}$ . En posant  $\varphi_0(x) = \ln(x + \lambda i)$  nous avons  $\varphi(x) = \varphi_0(x + s)$  et donc

$$\varphi'(x) = \varphi'_0(x + s) = \frac{1}{x + s + \lambda i} = \frac{1}{x + z}. \quad (26.152)$$

Tout va bien.

□

EXooAKEDooZgjocX

**Exemple 26.44.**

Un petit calcul d'intégrale, que nous avons déjà faite dans l'exemple 20.100 (avec la méthode de Rothstein-Trager). En passant par une décomposition en fractions simples :

$$\int \frac{1}{x^3 + x} = \int \left( \frac{1}{x} - \frac{1/2}{x - i} - \frac{1/2}{x + i} \right) \quad (26.153a)$$

$$= \ln(x) - \frac{1}{2} \ln(x - i) - \frac{1}{2} \ln(x + i) \quad (26.153b)$$

$$= \ln(x) - \frac{1}{2} \ln(x^2 + 1). \quad (26.153c)$$

SUBEQooRNQooScfSiG

Attention aux justifications. Il n'est pas vrai en général dans le cas de nombres complexes  $a$  et  $b$  que  $\ln(ab) = \ln(a) + \ln(b)$ . En effet, pour la partie réelle, ça passe parce que  $|ab| = |a||b|$ . Mais en ce qui concerne la partie imaginaire,

$$\arg(ab) \neq \arg(a) + \arg(b) \quad (26.154)$$

lorsque la somme dépasse les bornes de  $]-\pi, \pi]$ . Le passage à (26.153c) fonctionne parce que dans le cas particulier des nombres  $x + i$  et  $x - i$ , les arguments se somment à zéro :  $\arg(x + i) + \arg(x - i) = 0$ .  $\triangle$

**26.2.8 Logarithme sur un chemin**

DEFooBBGFooCEdsFR

**Définition 26.45** (logarithme continu).

Soient un espace topologique  $X$ , et une application  $f: X \rightarrow \mathbb{C}^*$ . Nous disons que  $g: X \rightarrow \mathbb{C}^*$  est un **logarithme** de  $f$  si pour tout  $x \in X$  nous avons

$$f(x) = \exp(g(x)). \quad (26.155)$$

DEFooOCDGooGyvvWi

**Définition 26.46** (Détermination du logarithme).

Soit un chemin<sup>15</sup>  $\gamma: [a, b] \rightarrow \mathbb{C}^*$ . Nous disons qu'une application  $g: \gamma([a, b]) \rightarrow \mathbb{C}^*$  est une **détermination du logarithme** sur  $\gamma$  si

$$\exp((g \circ \gamma)(t)) = \gamma(t) \quad (26.156)$$

pour tout  $t \in [a, b]$ .

THOooUPANooMiECqe

**Théorème 26.47.**

Tout chemin dans  $\mathbb{C}^*$  admet une détermination continue du logarithme<sup>16</sup>, et si  $l$  est une détermination sur le chemin  $\gamma$ , nous avons

$$\int_{\gamma} \frac{dz}{z} = l(\gamma(b)) - l(\gamma(a)). \quad (26.157)$$

Le théorème de Borsuk parle de lien entre logarithme continu et homotopie de fonctions à la fonction constante. Il s'applique donc bien de concert avec la proposition 26.16 qui dit que si  $f$  et  $g$  sont homotopes, alors  $f/g$  est homotope à une constante.

15. Définition 20.75.

16. Définition 26.46.

THOooTCUMooEByCKg

**Théorème 26.48** (Théorème de Borsuk<sup>[234]</sup>).

Soient un compact  $K$  de  $\mathbb{R}^n$  ainsi qu'une application continue  $f: K \rightarrow \mathbb{C}^*$ . Les propriétés suivantes sont équivalentes :

ITEMooKZYDooKoEEbl

(1)  $f$  admet un logarithme continu<sup>17</sup>.

ITEMooXVNXooVAHklr

(2)  $f$  est homotope à l'application constante  $u: K \rightarrow \mathbb{C}^*$ ,  $u(z) = 1$ .

ITEMooQDHXooObjxLA

(3)  $f$  admet une extension continue  $\tilde{f}: \mathbb{R}^n \rightarrow \mathbb{C}^*$ .

*Démonstration.* En plusieurs points.

(1) **(1)  $\Rightarrow$  (3)** Soit un logarithme continu  $g: K \rightarrow \mathbb{C}$  de  $f$ . Vu que  $K$  est fermé, le théorème de Tietze **26.22** dit que  $g$  possède une extension continue  $\tilde{g}: \mathbb{R}^n \rightarrow \mathbb{C}$ . Nous posons

$$\begin{aligned} \tilde{f}: \mathbb{R}^n &\rightarrow \mathbb{C}^* \\ x &\mapsto \exp(\tilde{g}(x)). \end{aligned} \quad (26.158)$$

L'application  $\tilde{f}$  est continue parce que  $\exp$  et  $\tilde{g}$  le sont. Elle est une extension de  $f$  parce que si  $x \in K$ , nous avons

$$\tilde{f}(x) = \exp(\tilde{g}(x)) = \exp(g(x)) = f(x) \quad (26.159)$$

parce que  $\tilde{g}(x) = g(x)$  et  $g$  est un inverse de  $\exp$ .

(2) **(3)  $\Rightarrow$  (2)** Soit une extension continue  $\tilde{f}: \mathbb{R}^n \rightarrow \mathbb{C}^*$  de  $f$ . Soit  $x_0 \in \mathbb{R}^n$ . Nous posons

$$\begin{aligned} H: [0, 1] \times K &\rightarrow \mathbb{C}^* \\ (t, x) &\mapsto \tilde{f}((1-t)x + tx_0). \end{aligned} \quad (26.160)$$

L'application  $H$  est continue, et pour  $x \in K$  elle vérifie  $H(0, x) = \tilde{f}(x) = f(x)$  ainsi que  $H(1, x) = \tilde{f}(x_0)$ .

Donc  $f$  est homotope à l'application constante  $\tilde{f}(x_0)$ . Vu que  $\mathbb{C}^*$  est connexe par arcs, toutes les applications constantes sont homotopes<sup>18</sup>. Donc  $f$  est homotope à  $\tilde{f}(x_0)$  qui est homotope à  $u$ . L'homotopie étant une relation d'équivalence<sup>19</sup>,  $f$  est homotope à  $u$ .

(3) **(2)  $\Rightarrow$  (1)** Nous considérons l'homotopie  $H: [0, 1] \times K \rightarrow \mathbb{C}^*$  entre  $f$  et  $u$ . En particulier pour tout  $x \in K$  nous avons  $H(0, x) = f(x)$  et  $H(1, x) = 1$ .

La partie  $[0, 1] \times K$  est compacte<sup>20</sup> et  $|H|$  y est continue. Donc elle atteint ses bornes. Mais elle prend ses valeurs dans  $\mathbb{C}^*$ ; donc

$$\inf_{x,t} |H(x, t)| > 0. \quad (26.161)$$

Et d'ailleurs cet infimum est un minimum. Nous considérons la norme suivante sur  $\mathbb{R} \times \mathbb{R}^n$  :

$$\|(t, x)\|_\infty = \max(|t|, |x|). \quad (26.162)$$

Toutes les normes étant équivalentes<sup>21</sup> sur  $\mathbb{R} \times \mathbb{R}^n$ , nous pouvons caractériser la compacité, la continuité et tout ça en termes de cette norme. L'application  $H$  est uniformément continue<sup>22</sup>.

Soit  $\epsilon > 0$ , il existe  $\eta > 0$  tel que si  $\|(t, x) - (s, y)\|_\infty < \eta$  alors  $\|H(t, x) - H(s, y)\| < \epsilon$ .

Soient un entier  $n > \frac{1}{\eta} + 1$ ,  $x \in K$  et  $k \in \{0, \dots, n\}$ . Nous avons

$$\left\| \left( \frac{k}{n}, x \right) - \left( \frac{k+1}{n}, x \right) \right\|_\infty = \left\| \left( -\frac{1}{n}, 0 \right) \right\| = \frac{1}{n} < \eta. \quad (26.163)$$

17. Définition **26.45**.

18. Lemme **26.15**.

19. Lemme **26.14**.

20. Théorème **7.367**.

21. Équivalence de normes, théorème **10.57**.

22. Théorème de Heine **12.80**.

Nous avons donc également

$$\left| H\left(\frac{k}{n}, x\right) - H\left(\frac{k+1}{n}, x\right) \right| < \epsilon. \quad (26.164)$$

Nous posons

$$\begin{aligned} F_k &: K \rightarrow \mathbb{C}^* \\ x &\mapsto H\left(\frac{k}{n}, x\right). \end{aligned} \quad (26.165)$$

Cette application est continue sur  $K$  et pour tout  $x \in K$  nous avons <sup>SUBEQSooLJIJooYkxyMO</sup>

$$\left| \frac{F_k(x)}{F_{k+1}(x)} - 1 \right| = \left| \frac{F_k(x) - F_{k+1}(x)}{F_{k+1}(x)} \right| \quad (26.166a)$$

$$= \left| \frac{H\left(\frac{k}{n}, x\right) - H\left(\frac{k+1}{n}, x\right)}{H\left(\frac{k+1}{n}, x\right)} \right| \quad (26.166b)$$

$$< \frac{\epsilon}{\left| H\left(\frac{k+1}{n}, x\right) \right|}. \quad (26.166c)$$

Pour rappel, pour tout  $\epsilon > 0$ , il existe un  $\eta > 0$  tel qu'en posant  $n > \frac{1}{\eta} + 1$  nous avons (26.166). Nous choisissons  $\epsilon$  tel que  $0 < \epsilon < \inf_{(t,x)} |H(t,x)|$ , de telle sorte à avoir

$$\left| \frac{F_k(x)}{F_{k+1}(x)} - 1 \right| < \frac{\epsilon}{\left| H\left(\frac{k+1}{n}, x\right) \right|} < 1. \quad (26.167)$$

La fonction  $\ln: \mathbb{C}^* \rightarrow \mathbb{C}^*$  est continue sur  $B(1,1)$  <sup>23</sup>. Donc la fonction

$$\begin{aligned} h_k &: K \rightarrow \mathbb{C}^* \\ x &\mapsto \ln\left(\frac{F_k(x)}{F_{k+1}(x)}\right) \end{aligned} \quad (26.168)$$

est continue. Nous posons

$$g = \sum_{k=0}^{n-1} \ln\left(\frac{F_k}{F_{k+1}}\right), \quad (26.169)$$

qui est également continue sur  $K$ . Cette fonction  $g$  est un logarithme continu de  $f$  sur  $K$  parce que <sup>24</sup>

$$\exp(g(x)) = \exp\left(\sum_{k=0}^{n-1} \ln\left(\frac{f_k(x)}{F_{k+1}(x)}\right)\right) \quad (26.170a)$$

$$= \prod_{k=0}^{n-1} \exp\left(\ln\left(\frac{F_k(x)}{F_{k+1}(x)}\right)\right) \quad (26.170b)$$

$$= \prod_{k=0}^{n-1} \frac{F_k(x)}{F_{k+1}(x)} \quad (26.170c)$$

$$= \frac{F_0(x)}{F_n(x)} \quad (26.170d)$$

$$= \frac{H(0, x)}{H(1, x)} \quad (26.170e)$$

$$= \frac{f(x)}{1} \quad (26.170f)$$

$$= f(x). \quad (26.170g)$$

23. Le logarithme est celui défini en 26.30 et sa continuité est le théorème 26.35.

24. Dans le calcul suivant, nous utilisons entre autres la formule  $\exp(a+b) = \exp(a)\exp(b)$  de la proposition 18.9(2).

□

CORooXOOZooUJMKxu

**Corolaire 26.49** ([234]).

Soit un compact convexe  $K$  de  $\mathbb{R}^n$ . Toute application continue  $K \rightarrow \mathbb{C}^*$  admet un logarithme continu.

*Démonstration.* Soit une application continue  $f: K \rightarrow \mathbb{C}^*$ . Soit  $x_0 \in K$ . Vu que  $K$  est convexe,  $f$  est homotope à la fonction constante  $f(x_0)$ . Par arc-connexité de  $\mathbb{C}^*$ , l'application  $f$  est alors homotope à la fonction constante 1.

Le théorème de Borsuk 26.48 nous dit alors que  $f$  admet un logarithme continu. □

LEMooJNPTooScfSvA

**Lemme 26.50.**

Si  $K$  est compact dans  $\mathbb{C}$ , alors la partie  $\mathbb{C} \setminus K$  possède exactement une composante connexe non bornée.

*Démonstration.* La partie  $K$  étant compacte, elle est bornée. Il existe donc  $r > 0$  tel que  $K \subset B(0, r)$ . Soit  $A = \mathbb{C} \setminus B(0, r)$ . Cela est une partie connexe de  $\mathbb{C}$  et est donc contenue dans une composante connexe de  $\mathbb{C} \setminus K$ . Il y a donc existence d'une composante connexe non bornée de  $\mathbb{C} \setminus K$ .

Pour l'unicité, les autres composantes connexes de  $\mathbb{C} \setminus K$  sont contenues dans  $B(0, r)$  et sont donc bornées. □

**26.2.9 Lacets, indice et homotopie**

DEFooLFBNooGlvJmp

**Proposition-Définition 26.51** ([699]).

Soit  $\gamma$  un chemin fermé<sup>25</sup> dans  $\mathbb{C}$  que nous supposons continu et  $C^1$  par morceaux<sup>26</sup>. L'**indice** de la courbe  $\gamma$  est la fonction

$$\begin{aligned} \text{Ind}_\gamma: \mathbb{C} \setminus \gamma &\rightarrow \mathbb{Z} \\ z &\mapsto \frac{1}{2\pi i} \int_\gamma \frac{d\omega}{\omega - z}. \end{aligned} \quad (26.171)$$

Alors :

ITEMooHYHMooWcaJxP

- (1) La fonction  $\text{Ind}_\gamma$  prend ses valeurs dans  $\mathbb{Z}$ .
- (2) Elle est continue sur chaque composante connexe de  $\mathbb{C} \setminus \gamma$ .
- (3) La fonction indice est constante sur chaque composante connexe<sup>27</sup> de  $\mathbb{C} \setminus \gamma$ .
- (4) Elle est nulle sur les composantes non bornées de  $\mathbb{C} \setminus \gamma$ <sup>28</sup>.

*Démonstration.* Point par point.

(1) **Prend ses valeurs dans  $\mathbb{Z}$** 

Explicitons l'intégrale sur  $\gamma: [a, b] \rightarrow \mathbb{C}$  :

$$\text{Ind}_\gamma(z_0) = \frac{1}{2\pi i} \int_\gamma \frac{dz}{z - z_0} = \frac{1}{2\pi i} \int_a^b \frac{\gamma'(t)}{\gamma(t) - z_0} dt. \quad (26.172)$$

Note. Comme  $\gamma$  n'est peut-être pas dérivable en certains points, en réalité nous n'avons pas une intégrale entre  $a$  et  $b$ , mais une somme d'intégrales entre  $a_i$  et  $b_i$  avec  $a_0 = a$  et  $b_n = b$ . Nous nommons  $S$  l'ensemble de ces points.

Nous allons calculer l'exponentielle de ce qui est dans l'intégrale en posant

$$\begin{aligned} \varphi: [a, b] &\rightarrow \mathbb{C} \\ u &\mapsto \exp \left( \int_a^u \frac{\gamma'(t)}{\gamma(t) - z_0} dt \right). \end{aligned} \quad (26.173)$$

25. Par abus de langage, nous désignerons par  $\gamma$  à la fois le chemin et son image.

26. Définition 20.37

27. Définition 7.104.

28. Elle a une seule telle composante par le lemme 26.50, mais ce n'est pas important ici.

Nous savons par le lien entre primitive et intégrale (proposition 14.287) que la dérivée de  $x \mapsto \int_0^x f$  est  $f$  elle-même. Nous pouvons donc facilement dériver  $\varphi$  :

$$\varphi'(u) = \frac{\gamma'(u)}{\gamma(u) - z_0} \exp \left( \int_a^u \frac{\gamma'(t)}{\gamma(t) - z_0} dt \right) = \frac{\gamma'(u)\varphi(u)}{\gamma(u) - z_0}. \quad (26.174)$$

Cette égalité est vraie pour tout  $u \in [a, b]$  sauf en un nombre fini de points (les points de  $S$ ). Nous considérons maintenant

$$\begin{aligned} f: [a, b] &\rightarrow \mathbb{C} \\ t &\mapsto \frac{\varphi(t)}{\gamma(t) - z_0}, \end{aligned} \quad (26.175)$$

et nous dérivons :

$$f'(t) = \frac{\varphi'(t)(\gamma(t) - z_0) - \varphi(t)\gamma'(t)}{(\gamma(t) - z_0)^2} \quad (26.176a)$$

$$= \frac{\varphi'(t)}{\gamma(t) - z_0} - \frac{\varphi(t)\gamma'(t)}{(\gamma(t) - z_0)^2}. \quad (26.176b)$$

Nous y injectons l'expression (26.174) de  $\varphi'$  :

$$f'(t) = \frac{\gamma'(t)\varphi(t)}{(\gamma(t) - z_0)^2} - \frac{\varphi(t)\gamma'(t)}{(\gamma(t) - z_0)^2} = 0. \quad (26.177)$$

Comme  $f$  est continue<sup>29</sup> sur  $[a, b]$  et de dérivée nulle partout sauf sur  $S$ , elle est constante sur  $[a, b]$ . Nous avons donc  $f(t) = f(a)$  pour tout  $t$ , c'est-à-dire

$$\frac{\varphi(t)}{\gamma(t) - z_0} = \frac{1}{\gamma(a) - z_0}. \quad (26.178)$$

Nous pouvons isoler de là une valeur pour  $\varphi(t)$  :

$$\varphi(t) = \frac{\gamma(t) - z_0}{\gamma(a) - z_0}. \quad (26.179)$$

C'est le moment de calculer  $\varphi(b)$ . Dans le calcul suivant, nous utilisons  $\gamma(a) = \gamma(b)$  :

$$\varphi(b) = \frac{\gamma(b) - z_0}{\gamma(a) - z_0} = \frac{\gamma(b) - z_0}{\gamma(b) - z_0} = 1. \quad (26.180)$$

Nous avons prouvé que

$$\exp \left( \int_a^b \frac{\gamma'(t)}{\gamma(t) - z_0} dt \right) = 1. \quad (26.181)$$

Pour qu'une exponentielle soit égale à 1, il faut que ce qui se trouve dedans soit égal à  $2ik\pi$  pour un certain  $k$ . Il existe  $k \in \mathbb{Z}$  tel que

$$\int_a^b \frac{\gamma'(t)}{\gamma(t) - z_0} dt = 2i\pi k. \quad (26.182)$$

Pour cette valeur de  $k$  nous avons  $\text{Ind}_\gamma(z_0) = k$ .

- (2) **Une borne pour  $|\gamma'|$**  L'application  $\gamma$  est  $C^1$  par morceaux<sup>30</sup>. La restriction  $\gamma': [a_i, b_i] \rightarrow \mathbb{C}$  est continue sur un compact et est donc bornée. Pour chaque  $i$  nous avons donc une borne  $M_i$  pour  $|\gamma'j|$ . En nommant  $M$  le maximum des ces bornes  $M_i$  nous avons  $|\gamma'(t)| < M$  pour tout  $t \in [a, b]$ .

29. Fait important : nous utilisons la continuité de  $f$ , pas de l'indice. Nous n'avons pas encore démontré que l'indice est une fonction continue.

30. Nous utilisons ici crucialement le fait que les morceaux sont fermés.



- (3) **L'indice est ponctuellement continu** Nous utilisons le théorème 17.15 avec l'application

$$f: [a, b] \times \mathbb{C} \setminus \gamma \rightarrow \mathbb{C}$$

$$(t, z) \mapsto \frac{\gamma'(t)}{\gamma(t) - z}. \quad (26.183)$$

Sauf que non. Il ne va pas être possible de majorer cette fonction uniformément en  $z$ , comme il le faudrait pour remplir la condition (3). La raison est que quand  $z$  s'approche de  $\gamma$ , le dénominateur n'est pas contrôlé.

Nous allons donc restreindre le domaine de  $f$ . Soit  $z_0 \in \mathbb{C} \setminus \gamma$ . Vu que  $\gamma$  est continue et que  $[a, b]$  est compact, la fonction  $t \mapsto 1/(\gamma(t) - z_0)$  atteint une borne inférieure. Disons que cette borne est atteinte pour  $t = t_0$ . Nous considérons  $r < d(z_0, \gamma(t_0))$  et nous prenons une borne  $M$  pour  $|\gamma'|$ . Nous avons alors, pour tout  $t$  que

$$\frac{|\gamma'(t)|}{|\gamma(t) - z_0|} < \frac{M}{r}. \quad (26.184)$$

Le  $z_0$  étant fixé, la fonction constante  $M/r$  majore  $f$  uniformément en  $t$ . La théorème de continuité sous l'intégrale 17.15 dit alors que l'indice est continue en  $z_0$ .

- (4) **Continuité globale** La fonction indice étant continue en chaque point de  $\mathbb{C} \setminus \gamma$ , elle est continue sur  $\mathbb{C} \setminus \gamma$  (théorème 7.262).
- (5) **Sur les composantes connexes** Soit une composante connexe  $C$  de  $\mathbb{C} \setminus \gamma$ . L'application  $\text{Ind}_\gamma$  est continue et à valeurs entières sur  $C$ . La proposition 7.270 dit alors que l'indice est constant sur  $C$ .
- (6) **Nulle sur la non bornée** En majorant  $|\gamma'|$  par  $M$ , nous avons

$$|\text{Ind}_\gamma(z_0)| \leq \frac{1}{2\pi} \int_a^b \frac{|\gamma'(t)|}{|\gamma(t) - z_0|} dt \leq \frac{M}{2\pi} \int_a^b \frac{dt}{|\gamma(t) - z_0|}. \quad (26.185)$$

Soit  $\epsilon > 0$  assez petit pour que  $\epsilon M(b-a)/2\pi < 1$ . Vu que  $\gamma$  est bornée, il existe  $R > 0$  tel que  $\gamma(t) \in B(0, R)$  pour tout  $t$ . Nous considérons  $r$  assez grand pour que  $r - R > 1/\epsilon$ .

Comme nous étudions  $\text{Ind}_\gamma$  sur une composante non bornée, nous considérons  $|z_0| > r$ . Avec ça nous avons  $|\gamma(t) - z_0| > r - R$  et donc

$$|\text{Ind}_\gamma(z_0)| \leq \frac{M}{2\pi} \int_a^b \epsilon dt = \frac{M(b-a)}{2\pi} \epsilon < 1. \quad (26.186)$$

La dernière inégalité est un choix de  $\epsilon$  fait au départ.

Étant donné que  $\text{Ind}_\gamma(z_0) \in \mathbb{Z}$ , nous avons forcément  $\text{Ind}_\gamma(z_0) = 0$ . Et comme l'indice est constant sur les composantes connexe, il est nul partout.

□

PROPooXWULooFZgHfL

### Proposition 26.52.

Soit  $p \in \mathbb{C}$ . Soit un chemin  $\gamma: [0, 1] \rightarrow \mathbb{C} \setminus \{p\}$ . Quelques hypothèses :

- (1) Il existe une demi-droite  $D$  d'origine  $p$  et tel que  $\gamma^{-1}(D)$  est fini.
- (2) Le chemin  $\gamma$  traverse  $D$  en tout point de  $\gamma^{-1}(D)$ .
- (3) Soit  $n^+$  le nombre de points de  $\gamma^{-1}(D)$  où  $\gamma$  traverse positivement  $D$  et  $n^-$  idem négativement.

Alors

$$\text{Ind}(\gamma, p) = n^+ - n^-. \quad (26.187)$$

PROPooEKFHooOWeIMk

### Proposition 26.53 ([234]).

Soient  $p \in \mathbb{C}$  ainsi que deux chemins  $\gamma_1, \gamma_2: [a, b] \rightarrow \mathbb{C} \setminus \{p\}$ . Si  $\gamma_1$  et  $\gamma_2$  sont homotopes dans  $\mathbb{C} \setminus \{p\}$ . Alors

$$\text{Ind}(\gamma_1, p) = \text{Ind}(\gamma_2, p). \quad (26.188)$$

*Démonstration.* Posons

$$\begin{aligned} f: \mathbb{C} \setminus \{p\} &\rightarrow \mathbb{C} \\ z &\mapsto \frac{1}{z-p}. \end{aligned} \quad (26.189)$$

La partie  $\mathbb{C} \setminus \{p\}$  est ouverte. Le théorème 26.18 d'invariance de l'intégrale par homotopie nous indique que

$$\text{Ind}(\gamma_1, p) = \frac{1}{2\pi i} \int_{\gamma_1} f = \frac{1}{2\pi i} \int_{\gamma_2} f = \text{Ind}(\gamma_2, p). \quad (26.190)$$

□

DefECnFJQp

### Définition 26.54.

Si  $\gamma_1$  et  $\gamma_2$  sont deux lacets en  $x_0 \in X$  (un espace topologique), une **équivalence d'homotopie** est une application  $f: [0, 1] \times [0, 1] \rightarrow X$  telle que

- (1)  $f(0, t) = \gamma_1(t)$  pour tout  $t$  ;
- (2)  $f(1, t) = \gamma_2(t)$  pour tout  $t$  ;
- (3) pour chaque  $t \in [0, 1]$ , l'application  $s \mapsto f(s, t)$  est continue ;
- (4) pour chaque  $s \in [0, 1]$ , l'application  $t \mapsto f(s, t)$  est un lacet basé en  $x_0$ .

LEMooMEICooOOGzdf

### Lemme 26.55.

Si  $\gamma$  est un cercle de centre  $z_0 \in \mathbb{C}$  et de rayon  $r$ , alors

$$\text{Ind}_\gamma(z) = \begin{cases} 1i & \text{si } z \in B(z_0, r) \\ 0 & \text{sinon.} \end{cases} \quad (26.191)$$

*Démonstration.* La seconde ligne provient directement de la proposition 26.51. Pour la première, le cercle  $\gamma$  se paramètre par

$$\gamma(\theta) = z_0 + re^{i\theta}. \quad (26.192)$$

Nous commençons par calculer  $\text{Ind}_\gamma(z_0)$ . Il vaut

$$\frac{1}{2\pi i} \int_\gamma \frac{d\omega}{\omega - z_0} = \frac{1}{2\pi i} \int_0^{2\pi} \frac{1}{re^{i\theta}} ire^{i\theta} d\theta = 1. \quad (26.193)$$

Nous considérons maintenant un point  $p \in B(z_0, r)$  et nous calculons  $\text{Ind}_\gamma(p)$ . D'abord si  $\sigma$  est le cercle  $\sigma(t) = \gamma(t) + p - z_0 = re^{it} + p$  de rayon  $r$  centré en  $p$ , nous avons

$$\text{Ind}_\sigma(p) = 1 \quad (26.194)$$

par ce que nous venons de voir.

Ensuite les chemins  $\sigma$  et  $\gamma$  sont homotopes dans  $\mathbb{C} \setminus \{p\}$ . En effet nous posons

$$H(u, t) = (1 - u)\gamma(t) + u\sigma(t). \quad (26.195)$$

Nous avons

$$H(u, t) - p = (1 - u)z_0 + re^{it} + (u - 1)p = (1 - u)(z_0 - p) + re^{it}. \quad (26.196)$$

En termes de modules, nous avons d'une part

$$|(1 - u)(z_0 - p)| \leq |z_0 - p| < r \quad (26.197)$$

et  $|re^{it}| = r$ . Donc la somme des deux ne peut pas faire zéro. Nous avons donc bien  $H(u, t) \neq p$  pour tout  $u \in [0, 1]$  et pour tout  $t \in [0, 2\pi]$ .

Les chemins  $\gamma$  et  $\sigma$  étant homotopes dans  $\mathbb{C} \setminus \{p\}$ , la proposition 26.53 dit que

$$\text{Ind}_\gamma(p) = \text{Ind}_\sigma(p) = 1. \quad (26.198)$$

□

PROPOGAOIooFTOuli

**Proposition 26.56** ([234]).

Soit un intervalle  $J \subset \mathbb{R}$ . Nous considérons l'application

$$\begin{aligned}\psi: \mathcal{L}(J, \mathbb{C}^*) &\rightarrow \mathbb{Z} \\ \gamma &\mapsto \text{Ind}(\gamma, 0)\end{aligned}\tag{26.199}$$

où  $\mathcal{L}(J, \mathbb{C}^*)$  est l'ensemble des lacets  $J \rightarrow \mathbb{C}^*$  sur lequel nous considérons la loi de groupe multiplicatif tandis que nous considérons la loi additive sur  $\mathbb{Z}$ . Nous avons :

ITEMooKCRlooSEyhlp

- (1)  $\psi$  est effectivement à valeurs dans  $\mathbb{Z}$ ,
- (2)  $\psi$  est un morphisme de groupes surjectif.
- (3) Le noyau de  $\psi$  est

ITEMooZFXTooDCXTVU

$$\ker(\psi) = \{\text{lacets homotopes dans } \mathbb{C}^* \text{ à un lacet constant}\}.\tag{26.200}$$

*Démonstration.* Pour (1), c'est la proposition 26.51(1).

Soient deux lacets  $\gamma_1$  et  $\gamma_2$  dans  $\mathbb{C}^*$ . Le théorème 26.47 nous permet de considérer des logarithmes continus  $l_i: [a, b] \rightarrow \mathbb{C}$  le long de  $\gamma_i$ . L'application  $l_1 + l_2$  est une détermination du logarithme le long de  $\gamma_1\gamma_2$  parce que<sup>31</sup>

$$\exp(l_1(t) + l_2(t)) = \exp(l_1(t)) \exp(l_2(t)) = \gamma_1(t)\gamma_2(t).\tag{26.201}$$

- (1) **Morphisme** Nous avons

$$\text{Ind}(\gamma_1\gamma_2, 0) = \frac{1}{2\pi i} \int_a^b f((\gamma_1\gamma_2)(t))(\gamma_1\gamma_2)'(t)dt\tag{26.202a}$$

$$= \frac{1}{2\pi i} \int_{\gamma_1\gamma_2} \frac{dz}{z}\tag{26.202b}$$

$$= (l_1 + l_2)(b) - (l_1 + l_2)(a)\tag{26.202c}$$

$$= (l_1(b) - l_1(a)) + (l_2(b) - l_2(a))\tag{26.202d}$$

$$= \text{Ind}(\gamma_1, 0) + \text{Ind}(\gamma_2, 0).\tag{26.202e}$$

Nous avons prouvé que  $\psi(\gamma_1\gamma_2) = \psi(\gamma_1) + \psi(\gamma_2)$ , et donc que  $\psi$  est un morphisme.

- (2) **Surjection** Nous considérons le lacet

$$\begin{aligned}\gamma: [a, b] &\rightarrow \mathbb{C}^* \\ t &\mapsto \exp(2i\pi kt(b-a)),\end{aligned}\tag{26.203}$$

et nous montrons que  $\psi(\gamma) = k$ . D'abord nous remarquons que

$$\gamma'(t) = \gamma(t) \frac{2i\pi k}{(b-a)},\tag{26.204}$$

et ensuite nous calculons :

$$\psi(\gamma) = \frac{1}{2\pi i} \int_a^b \frac{1}{\gamma(t)} \gamma'(t) dt\tag{26.205a}$$

$$= \frac{1}{2\pi i} \int_a^b \gamma(t) \frac{2i\pi k}{(b-a)} \gamma(t) dt\tag{26.205b}$$

$$= \frac{1}{2\pi i} \int_a^b \frac{2i\pi k}{(b-a)} dt\tag{26.205c}$$

$$= k.\tag{26.205d}$$

31. Nous ne nous lasserons jamais de citer la proposition 18.9(2).

(3) **Noyau** Notons pour la simplicité

$$\mathcal{H} = \{\text{lacets homotopes dans } \mathbb{C}^* \text{ à un lacet constant}\}. \quad (26.206)$$

Nous notons  $u$  le chemin constant  $u(t) = 1$ . Nous avons  $\psi(u) = 0$  parce que, en notant  $f(z) = 1/z$ ,

$$\text{Ind}(u, 0) = \frac{1}{2\pi i} \int_u f(z) dz = \frac{1}{2\pi i} \int_0^1 f(u(t)) u'(t) dt = 0 \quad (26.207)$$

parce que  $u'(t) = 0$  pour tout  $t$ . Nous avons donc  $u \in \ker(\psi)$ . Si  $\gamma \in \mathcal{H}$ , alors, utilisant le théorème de Cauchy homotopique 26.18(2), nous avons  $\text{Ind}(\gamma, 0) = \text{Ind}(u, 0) = 0$ . Donc  $\mathcal{H} \subset \ker(\psi)$ .

Soient  $\gamma \in \ker(\psi)$  et une détermination continue  $l$  du logarithme le long de  $\gamma$ . Nous avons

$$0 = \psi(\gamma) = l(b) - l(a). \quad (26.208)$$

Nous posons

$$\begin{aligned} H: [0, 1] \times [a, b] &\rightarrow \mathbb{C}^* \\ (s, t) &\mapsto \exp(sl(a) + (1-s)l(t)). \end{aligned} \quad (26.209)$$

Cette application continue vérifie  $H(0, t) = e^{l(t)} = \gamma(t)$  et  $H(1, t) = e^{l(a)} = \gamma(a)$ . Donc elle est une homotopie entre  $\gamma$  et le chemin constant  $\gamma(a)$ . Encore par Cauchy homotopique,  $\psi(\gamma) = 0$  parce que  $\psi$  est nul sur les chemins constants.

□

LEM00CTIB00NRAyZH

**Lemme 26.57** ([1]).

Si  $\gamma$  est un lacet dans  $\mathbb{C} \setminus \{p\}$ , alors

$$\text{Ind}(\gamma, p) = \text{Ind}(\gamma - p, 0) \quad (26.210)$$

PROP00PRII00KWCHBZ

**Proposition 26.58** ([234]).

Soient  $p \in \mathbb{C}$  ainsi que deux lacets  $\gamma_1$  et  $\gamma_2$  dans  $\mathbb{C} \setminus \{p\}$ . Si

$$\text{Ind}(\gamma_1, p) = \text{Ind}(\gamma_2, p), \quad (26.211)$$

alors  $\gamma_1$  et  $\gamma_2$  sont homotopes dans  $\mathbb{C} \setminus \{p\}$ .

*Démonstration.* Vu que  $\gamma_1$  ne passe pas par  $p$ , nous pouvons considérer le chemin

$$\begin{aligned} \gamma: [a, b] &\rightarrow \mathbb{C}^* \\ t &\mapsto \frac{\gamma_0(t) - p}{\gamma_1(t) - p}. \end{aligned} \quad (26.212)$$

En posant  $\sigma_i(t) = \gamma_i(t) - p$ , nous avons

$$\text{Ind}(\gamma, 0) = \text{Ind}(\sigma_0 \sigma_1^{-1}, 0) \quad (26.213a)$$

$$= \text{Ind}(\sigma_0, 0) + \text{Ind}(\sigma_1^{-1}, 0) \quad \text{prop. 26.56} \quad (26.213b)$$

$$= \text{Ind}(\sigma_0, 0) - \text{Ind}(\sigma_1, 0) \quad (26.213c)$$

$$= \text{Ind}(\gamma_0, p) - \text{Ind}(\gamma_1, p) \quad \text{lem. 26.57} \quad (26.213d)$$

$$= 0. \quad (26.213e)$$

Vu que l'indice de  $\gamma$  est nul, ce chemin est homotope dans  $\mathbb{C}^*$  au chemin constant  $u$  par la proposition 26.56(3). Soit une homotopie  $H: [0, 1] \times [a, b] \rightarrow \mathbb{C}^*$  entre  $\gamma$  et  $u$ . Et nous posons

$$\begin{aligned} S: [0, 1] \times [a, b] &\rightarrow \mathbb{C}^* \\ (s, t) &\mapsto (\gamma_1(t) - p)H(s, t). \end{aligned} \quad (26.214)$$

Cette application est une homotopie entre

$$t \mapsto S(0, t) = (\gamma_2(t) - p)H(0, t) = (\gamma_2(t) - p)\gamma(t) = \gamma_1(t) - p \quad (26.215)$$

et

$$t \mapsto S(1, t) = (\gamma_2(t) - p)H(1, t) = (\gamma_2(t) - p)u(t) = \gamma_2(t) - p. \quad (26.216)$$

Donc  $S + p$  est une homotopie entre  $\gamma_1$  et  $\gamma_2$  dans  $\mathbb{C} \setminus \{p\}$ .  $\square$

PROPooCFM FooXjlfV

**Proposition 26.59** ([234]).

Soit un espace topologique  $X$  localement connexe par arcs. Soit une application continue  $f: X \rightarrow \mathbb{C}^*$ . Il y a équivalence entre

- (1)  $f$  admet un logarithme continu.
- (2) Pour tout lacet  $\gamma$  dans  $X$ , nous avons  $\text{Ind}(f \circ \gamma, 0) = 0$ .

### 26.2.10 Théorème de Cauchy et analyticit 

Cette sous-section veut prouver le th or me de Cauchy.

ThoUHztQe

**Th or me 26.60** (Formule de Cauchy).

Soient  $\Omega$  ouvert dans  $\mathbb{C}$ ,  $z_0 \in \Omega$  et  $f$  une fonction holomorphe sur  $\Omega$ . Soit  $r > 0$  tel que  $B(z_0, r) \subset \Omega$ . Alors pour tout  $z \in B(z_0, r)$  nous avons

$$f(z) = \frac{1}{2\pi i} \int_{\partial B(z_0, r)} \frac{f(\omega)}{\omega - z} d\omega. \quad (26.217)$$

EaPzUABM

*D monstration.* Soit  $z \in B(z_0, r)$ . Consid rons la fonction

$$g(\omega) = \begin{cases} \frac{f(\omega) - f(z)}{\omega - z} & \text{si } \omega \neq z \\ f'(z) & \text{si } \omega = z. \end{cases} \quad (26.218)$$

Cette fonction est holomorphe sur  $B(z_0, r) \setminus \{z\}$  et continue en  $z$ . Elle v rifie donc la proposition 26.11 et nous avons

$$\int_{\gamma} g = 0 \quad (26.219)$$

o   $\gamma$  est le cercle de centre  $z_0$  et de rayon  $r$ . Nous avons donc

$$0 = \int_{\gamma} \frac{f(\omega)}{\omega - z} - \int_{\gamma} \frac{f(z)}{\omega - z}, \quad (26.220)$$

et ayant d j  calcul  la seconde int grale dans l'exemple 26.55 nous en d duisons

$$\int_{\gamma} \frac{f(\omega)}{\omega - z} d\omega = 2\pi i f(z), \quad (26.221)$$

ce qu'il fallait.  $\square$

ThomePOdd

**Th or me 26.61.**

Soient  $\Omega$  ouvert dans  $\mathbb{C}$  et  $f$  holomorphe sur  $\Omega$ . Soient encore  $z_0 \in \Omega$  et  $r_0$  tels que  $B(z_0, r_0) \subset \Omega$ . Alors :

ITEMooYWSOooHJtxGr

- (1) Sur  $B(z_0, r_0)$ , la fonction  $f$  s' crit

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \quad (26.222)$$

(2) Nous avons

$$a_n = \frac{f^{(n)}(z_0)}{n!} = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\omega)}{(\omega - z_0)^{n+1}} d\omega \quad (26.223)$$

où  $\gamma = \partial B(z_0, r)$  avec  $|z - z_0| < r < r_0$ .

ItemMRRTTooMChmuZ

(3) En particulier  $f$  est infiniment dérivable.

*Démonstration.* Soit  $r > 0$  tel que  $|z - z_0| < r < r_0$ . La formule de Cauchy (théorème 26.60) nous dit que

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\omega)}{\omega - z} d\omega \quad (26.224)$$

où  $\gamma = \partial B(z_0, r)$ . Nous pouvons paramétrer ce chemin par  $\omega = z_0 + re^{i\theta}$  et  $\theta \in [0, 2\pi]$ . Nous avons

$$f(z) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{z_0 + re^{i\theta} - z} rie^{i\theta} d\theta \quad (26.225a)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{1 - e^{-i\theta}(z - z_0)/r} d\theta. \quad (26.225b)$$

Nous pouvons développer l'intégrande en puissance de  $(z - z_0)$  en utilisant la formule 15.507. Ici le rôle de  $x$  est tenu par

$$e^{-i\theta}(z - z_0)/r \quad (26.226)$$

dont le module est bien plus petit que 1, par hypothèse sur  $r$ . Nous avons donc

$$f(z) = \frac{1}{2\pi} \int_0^{2\pi} \sum_{n=0}^{\infty} f(z_0 + re^{i\theta}) e^{-in\theta} r^{-n} (z - z_0)^n d\theta. \quad (26.227)$$

L'art est maintenant de permuter la somme et l'intégrale. Pour cela nous remarquons que ce qui se trouve dans la somme est majoré en module par

$$M \left| \frac{z - z_0}{r} \right|^n \quad (26.228)$$

EgbykTLD

où  $M$  est le maximum de  $|f|$  sur  $\gamma$ . La borne (26.228) ne dépend pas de  $\theta$ ; par conséquent la convergence de la somme est uniforme en  $\theta$  par le critère de Weierstrass (théorème 12.396). Le théorème 15.1 s'applique<sup>32</sup> et nous pouvons permuter la somme avec l'intégrale.

Ce que nous trouvons est que

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad (26.229)$$

où

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) e^{-in\theta} r^{-n} d\theta = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\omega)}{(\omega - z_0)^{n+1}} d\omega. \quad (26.230)$$

Cette formule est valable pour  $|z - z_0| < r$ . Sur cette boule, la fonction est donc une série entière. La proposition 12.478 nous permet donc d'affirmer que  $f$  est partout infiniment continument dérivable (parce que en chaque point on a un voisinage sur lequel c'est vrai), et d'identifier les coefficients (qui, eux, ne sont valables que localement) sous la forme

$$a_n = \frac{f^{(n)}(z_0)}{n!}. \quad (26.231)$$

□

32. Étant donné que nous savions déjà que la somme était une fonction intégrable, nous sommes loin d'avoir utilisé toute la puissance du théorème.

CORooJISDooFgwOPh

**Corolaire 26.62.**

Soit une application  $f: B_{\mathbb{C}}(0, R) \rightarrow \mathbb{C}$  ainsi que des  $(a_n)$  dans  $\mathbb{C}$  tels que

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad (26.232)$$

EQooWLEQooOslGfP

pour tout  $z \in B(0, r)$ .

Alors  $a_0 = f(0)$  et  $a_1 = f'(0)$ .

*Démonstration.* En posant  $z = 0$  à droite dans (26.232) nous trouvons tout de suite  $f(0) = a_0$ . En ce qui concerne  $a_1$ , nous passons par la proposition 12.478<sup>33</sup> nous dit que sur  $B(0, r)$ , l'application  $f$  est de classe  $C^\infty$  et que

$$df_z(u) = \sum_{n=1}^{\infty} n a_n z^{n-1} u. \quad (26.233)$$

Notez que dans le théorème 12.477, il est dit  $\alpha_n(x - x_0)^{n-1}$  qui est l'application linéaire qui à  $u \in \mathbb{C}$  fait correspondre  $\alpha_n(x - x_0, \dots, x - x_0, u)$ . Dans notre cas, ce qui joue le rôle de  $\alpha_n$  est l'application  $(z_1, \dots, z_n) \mapsto a_n z_1, \dots, z_n$ .

Ce qui nous dit la relation (12.911), c'est que, si  $f$  est différentiable<sup>34</sup>, le lien entre dérivée et différentielle est que la différentielle est l'application qui consiste à multiplier par la dérivée. Bref, nous avons

$$f'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1}. \quad (26.234)$$

En posant  $z = 0$  nous trouvons bien  $f'(0) = a_1$ . □

CorwfHtJu

**Corolaire 26.63.**

Soit  $f$  une fonction continue sur un ouvert  $\Omega$  telle que pour toute boule  $B(a, r)$  contenue dans  $\Omega$ , nous ayons

$$f(a) = \frac{1}{2\pi i} \int_{\partial B(a, r)} \frac{f(\xi)}{\xi - a} d\xi. \quad (26.235)$$

Alors  $f$  est holomorphe.

*Démonstration.* Il suffit de recopier la démonstration du théorème 26.61 pour savoir que  $f$  se développe en série de puissances et est donc en particulier dérivable. □

Le fait qu'une fonction holomorphe soit  $C^\infty$  comme dit dans la proposition 26.61 permet de démontrer un résultat de dérivation sous l'intégrale, qui dépend de pouvoir majorer la différentielle.

PROPooZCLYooUaSMWA

**Proposition 26.64.**

Soit une fonction continue  $g: \mathbb{R} \times \mathbb{C} \rightarrow \mathbb{C}$ . Nous supposons que pour tout  $t$ , la fonction  $z \mapsto g(t, z)$  est  $\mathbb{C}$ -dérivable (définition 12.330) et différentiable. Soit  $B$  compact dans  $\mathbb{R}$  et la fonction

$$G(z) = \int_B g(t, z) dt. \quad (26.236)$$

que nous supposons exister pour tout  $z$ .

Alors

$$G'(z) = \int_B g'(t, z) dt \quad (26.237)$$

où le prime réfère à la  $\mathbb{C}$ -dérivée par rapport à la variable  $z$  à  $t$  fixé.

33. Qui infiniment plus général que ce dont nous avons besoin, mais qui fait quand même bien le boulot.

34. C'est le cas de notre  $f$ , comme nous venons de le dire.

*Démonstration.* Nous fixons  $z \in \mathbb{C}$  et nous considérons la suite de fonctions

$$g_i(t) = \frac{g(t, z + \epsilon_i) - g(t, z)}{\epsilon_i} \quad (26.238)$$

où  $\epsilon_i$  est une suite de nombres complexes tendant vers zéro ( $\epsilon_i \xrightarrow{\mathbb{C}} 0$ ). Si la limite existe et ne dépend pas de la suite choisie, alors  $\lim_{i \rightarrow \infty} g_i(t) = g'(t, z)$ . Et vu que  $g$  est supposée dérivable, c'est le cas.

Nous avons aussi, par linéarité de l'intégrale :

$$G'(z) = \lim_{i \rightarrow \infty} \int_B g_i(t) dt. \quad (26.239)$$

La difficulté est de permuter la limite et l'intégrale. Pour cela nous allons utiliser la convergence dominée de Lebesgue (théorème 14.207). Afin de majorer  $|g_i(t)|$  par une fonction intégrable en  $t$  (uniformément en  $i$ ), nous exploitons le théorème des accroissements finis, théorème 10.324. En notant  $dg$  la différentielle de  $g$  par rapport à  $z$  à  $t$  fixé, pour chaque  $t$  et chaque  $i$  nous avons

$$|g(t, z + \epsilon_i) - g(t, z)| \leq \sup_{\xi \in [z, z + \epsilon_i]} \|dg_\xi\| \|\epsilon_i\|. \quad (26.240)$$

Vu que  $z$  est fixé et que  $\xi$  est dans le compact  $[z, z + \epsilon_i]$  et que  $dg$  est continue (parce que la  $\mathbb{C}$ -dérivabilité implique la continuité de la différentielle parce que nous avons l'analyticité par le théorème 26.61), nous pouvons majorer  $\|dg_\xi\|$  par une constante  $M_i(z)$  qui dépend à priori de  $i$  et de  $z$ .

Heureusement, nous pouvons prendre a fortiori le supremum sur  $\overline{B(z, |\epsilon_i|)}$  (qui est tout autant compact) et supposer que  $|\epsilon_i|$  est strictement décroissante ; de toutes façons, il y a un maximum parce que  $|\epsilon_i| \rightarrow 0$ . Dans ce cas, il suffit de prendre le supremum de  $\|dg_\xi\|$  pour  $\xi \in \overline{B(z, |\epsilon_1|)}$  et ça contente tout le monde.

Quoi qu'il en soit nous avons une constante  $M(z)$  telle que

$$|g(t, z + \epsilon_i) - g(t, z)| \leq M(z) \|\epsilon_i\| \quad (26.241)$$

et donc  $|g_i(t)| \leq M(z)$ . La constante (par rapport à  $t$ )  $M(z)$  est évidemment intégrable sur le compact  $B$  et nous pouvons permuter la limite avec l'intégrale :

$$G'(z) = \lim_{i \rightarrow \infty} \int_B g_i(t) dt = \int_B \lim_{i \rightarrow \infty} g_i(t) dt = \int_B g'(t, z) dt. \quad (26.242)$$

□

PropZOkfmO

### Proposition 26.65.

Une fonction continue  $f$  est holomorphe si et seulement si la 1-forme différentielle  $f(z)dz$  est localement exacte.

*Démonstration.* Si  $f$  est holomorphe, alors nous avons vu que  $f$  était différentiable et que  $df_z = f(z)dz$  par la formule 12.923.

Dans le sens inverse, supposons que  $f(z)dz$  est localement exacte, et soit  $F$  telle que  $dF = f(z)dz$ . Ce que nous allons faire est montrer que la dérivée de  $F$  existe et vaut  $f$ . En effet, la définition de la différentielle nous dit que

$$\lim_{h \rightarrow 0} \left| \frac{F(z + h) - F(z) - dF_z(h)}{h} \right| = 0. \quad (26.243)$$

La limite vaut évidemment encore zéro si nous enlevons les modules :

$$0 = \lim_{h \rightarrow 0} \frac{F(z + h) - F(z) - f(z)h}{h} \quad (26.244a)$$

$$= \lim_{h \rightarrow 0} \frac{F(z + h) - F(z)}{h} - f(z). \quad (26.244b)$$



Donc  $F' = f$ . Cela montre que  $F$  est  $\mathbb{C}$ -dérivable et donc holomorphe. En conséquence du théorème 26.61, la fonction  $F$  est infiniment dérivable et  $f$  l'est alors aussi. La fonction  $f$  est donc holomorphe<sup>35</sup>.  $\square$

### 26.2.11 Théorème de Brouwer en dimension 2

Pour d'autres versions du théorème de Brouwer, voir la sous-section 20.5.1.

ThoLVViheK

**Théorème 26.66** (Brouwer en dimension 2[126]).

Soit  $\mathcal{B} = \overline{B(0,1)}$  la boule unité fermée de  $\mathbb{R}^2$ . Alors toute application continue de  $\mathcal{B}$  dans elle-même admet un point fixe.

*Démonstration.* Supposons que la fonction  $f \in C^0(\mathcal{B}, \mathcal{B})$  n'admette pas de points fixes sur  $\mathcal{B} = \overline{B(0,1)}$ . Pour  $x \in \mathcal{B}$  nous notons  $g(x)$  l'intersection entre  $\partial\mathcal{B}$  et la demi-droite allant de  $f(x)$  vers  $x$ . C'est bien parce que  $f$  n'a pas de points fixes que  $g$  est bien définie.

En reprenant le même début de la preuve de la proposition 20.30 nous savons que la fonction

$$\begin{aligned} g: \overline{B(0,1)} &\rightarrow \partial B(0,1) \\ x &\mapsto \lambda(x)(x - f(x)) + f(x) \end{aligned} \quad (26.245)$$

est continue. De plus  $g(x) = x$  sur  $\partial B(0,1)$ . Nous allons montrer qu'une telle fonction<sup>36</sup> ne peut pas exister.

Pour  $s \in [0,1]$  nous paramétrons le cercle  $\partial B(0,s)$  par

$$\begin{aligned} x_s: [0,1] &\rightarrow \partial B(0,s) \\ t &\mapsto (s \cos(2\pi t), s \sin(2\pi t)). \end{aligned} \quad (26.246)$$

Ensuite nous considérons les chemins

$$\begin{aligned} \gamma_s: [0,1] &\rightarrow \partial B(0,1) \\ t &\mapsto g \circ x_s(t). \end{aligned} \quad (26.247)$$

L'application  $\gamma_s$  est continue et  $\gamma_s(0) = \gamma_s(1)$ . Les chemins  $\gamma_s$  sont des lacets ; nous nous intéressons maintenant à l'indice au point 0 de  $\gamma_0$  et  $\gamma_1$ . D'une part  $\gamma_0(t) = g(0)$  (lacet constant) et  $\gamma_1(t) = e^{2i\pi t}$  (parce que  $g(x) = x$  sur le bord). Nous avons donc, en utilisant l'indice de la définition 26.51,

$$\text{Ind}_{\gamma_0}(0) = \frac{1}{2\pi i} \int_{\gamma_0} \frac{d\omega}{\omega} = \frac{1}{2\pi i} \int_0^1 \frac{\gamma'_0(t)}{\gamma_0(t)} dt = 0, \quad (26.248)$$

alors que

$$\text{Ind}_{\gamma_1}(0) = \frac{1}{2\pi i} \int_0^1 \frac{2i\pi e^{2i\pi t}}{e^{2i\pi t}} dt = 1. \quad (26.249)$$

Nous considérons l'homotopie

$$\begin{aligned} \gamma: [0,1] \times [0,1] &\rightarrow \overline{B(0,1)} \\ (s,t) &\mapsto \gamma_s(t) = (g \circ x_s)(t). \end{aligned} \quad (26.250)$$

Nous avons  $g(0) \neq 0$  parce que  $g$  prend ses valeurs sur le bord. Vu que c'est une équivalence d'homotopie<sup>37</sup> entre  $\gamma_1$  et  $\gamma_2$ , les indices devraient être égaux par le corolaire 26.19.  $\square$

35. Dire que la dérivée d'une fonction holomorphe est holomorphe est un raisonnement classique.

36. Qui est nommée *rétraction* de la sphère sur elle-même.

37. Définition 26.54

### 26.2.12 Principe des zéros isolés

LEMooYYZQooCImOgG

#### Lemme 26.67.

Si  $f$  est une fonction holomorphe<sup>38</sup> sur le compact  $K$ , alors il existe un polynôme  $P_f$  et une fonction holomorphe  $h_f$  ne s'annulant pas sur  $K$  telles que  $f = h_f P_f$ .

*Démonstration.* Soit une fonction  $f$  vérifiant les conditions. Si  $f$  est identiquement nulle, alors il suffit de prendre  $P_f = 0$  et c'est fait. Nous supposons donc que  $f$  n'est pas identiquement nulle.

#### (1) Quantité finie de racines

D'abord  $f$  ne peut s'annuler qu'un nombre fini de fois sur  $K$ . Sinon, on pourrait considérer une suite des racines<sup>39</sup> de  $f$  dans  $K$ . Vu qu'une suite dans un compact contient une sous-suite convergente (théorème 7.194), la fonction  $f$  aurait un point d'accumulation de racines. Alors le principe des zéros isolés (théorème 17.136) nous donne un ouvert sur lequel  $f$  est nulle et donc le corolaire 17.137 nous dit que  $f$  est identiquement nulle.

#### (2) Autour d'une racine

Bref, la fonction  $f$  possède un nombre fini de racines sur  $K$ . Soit  $z_0$  l'une d'elles.

Par le théorème 26.61(1), nous avons, sur un voisinage de  $z_0$  :

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \quad (26.251)$$

En particulier,  $0 = f(z_0) = a_0$ . Donc  $a_0 = 0$ . Soit  $k$ , le plus petit naturel pour lequel  $a_k \neq 0$ . Nous avons

$$f(z) = (z - z_0)^k g(z) \quad (26.252)$$

avec  $g(z) = \sum_{n=0}^{\infty} a_{k+n} (z - z_0)^n$ . Vu que  $a_k \neq 0$  nous avons  $g(z_0) \neq 0$ . Montrons à présent que  $g$  est holomorphe sur un voisinage de  $z_0$ . Vu que la série définissant  $g$  est une sous-série d'une série convergente sur un voisinage, elle converge sur un voisinage et la proposition 26.12 nous dit que  $g$  est  $\mathbb{C}$ -dérivable. C'est-à-dire holomorphe par définition.

#### (3) Autour de toutes les racines

Soient  $(z_i)$  les racines (en nombre fini). Pour chaque  $i$  nous avons une boule  $B(z_i, r_i)$  sur laquelle  $f = P_i g_i$  où  $P_i$  est un polynôme de la forme  $(z - z_i)^k$  et  $g_i$  est holomorphe sur  $B(z_i, r_i)$ . Nous définissons la fonction suivante :

$$h(z) = \begin{cases} \frac{f(z)}{\prod_k P_k(z)} & \text{si } z \neq z_i \\ \frac{g_i}{\prod_{k \neq i} P_k(z)} & \text{si } z = z_i. \end{cases} \quad (26.253)$$

Cette fonction ne s'annule jamais. Mais est-elle holomorphe ?

Si  $z \neq z_i$  (sous-entendu : pour tout  $i$ ), alors sur un voisinage,  $h = f / \prod P_k$  qui est un quotient de fonctions holomorphes dont le dénominateur ne s'annule pas. Elle est donc holomorphe sur ce voisinage par le lemme 12.339.

Pour les autres notons que pour tout  $z \in B(z_i, r_i)$ ,

$$h = \frac{g_i}{\prod_{k \neq i} P_k}. \quad (26.254)$$

Cela est encore un quotient dont le dénominateur ne s'annule pas<sup>40</sup>.

#### (4) La réponse

Nous avons, pour tout  $z \in K$  :

$$f(z) = h(z) \prod_k P_k(z). \quad (26.255)$$

38. Définition 12.332.

39. Notez l'utilisation de la proposition 1.201 que je vous invite à ne pas considérer comme une trivialité absolue.

40. Nous avons choisi les  $r_i$  de telle sorte que les boules ne s'intersectent pas.

□

Afin de détendre l'atmosphère, nous allons laisser tomber l'analyse quelques instants et prouver un résultat d'algèbre.

PROPOVWRPOOGQMenV

**Proposition 26.68** ([1, 700]).

L'anneau des fonctions holomorphes sur un compact<sup>41</sup> donné de  $\mathbb{C}$  est principal<sup>42</sup>.

*Démonstration.* Nous nommons  $A$  l'ensemble des fonctions holomorphes sur le compact  $K$ , et  $J$  un idéal de  $A$ .

(1)  **$A$  est un anneau**

Le point délicat de la définition 1.90 est le fait que la somme et le produit d'éléments de  $A$  sont des éléments de  $A$  parce que les résultats type « la somme de deux fonctions holomorphes est holomorphe » sont valides sur des ouverts alors que nous sommes ici sur un compact. Soient  $f$  et  $g$  dans  $A$  ; nous nommons  $\Omega_f$  et  $\Omega_g$  des ouverts contenant  $K$  tels que  $f$  est holomorphe sur  $\Omega_f$  et  $g$  sur  $\Omega_g$ .

L'ensemble  $\Omega_f \cap \Omega_g$  est un ouvert (intersection d'ouverts) contenant  $K$  et sur lequel  $f$  et  $g$  sont holomorphes. Donc  $f + g$  et  $fg$  y sont holomorphes.

(2) **Engendré par des polynômes**

Pour chaque  $f \in J$  nous écrivons  $f = P_f h_f$  en vertu de la décomposition donnée par le lemme 26.67. Vu que  $h_f$  ne s'annule pas,  $1/h_f$  est encore holomorphe sur  $K$  et nous déduisons que  $P_f = f/h_f$  est dans  $J$ . La partie

$$S = \{P_f \text{ tel que } f \in J\} \quad (26.256)$$

est génératrice de  $J$  parce que, par construction, tous les éléments de  $J$  sont des produits d'éléments de  $S$  par des fonctions holomorphes sur  $K$  (donc, des éléments de  $A$ ). Mais tous les éléments de  $S$  sont dans  $J$ , donc  $(S) = J$ .

(3) **Un polynôme pour tous les engendrer**

Soit  $M$ , l'idéal de  $\mathbb{C}[X]$  engendré par  $S$ . Attention :  $J$  est l'idéal de  $A$  engendré par  $S$ . Mais l'idéal de  $\mathbb{C}[X]$  engendré par  $S$  est peut-être autre chose. Vu que  $\mathbb{C}$  est un corps, le lemme 3.98 dit que  $\mathbb{C}[X]$  est principal. Donc  $M$  est un idéal principal de  $\mathbb{C}[X]$  et nous avons un polynôme  $p \in \mathbb{C}[X]$  tel que

$$M = \mathbb{C}[X]p. \quad (26.257)$$

Si vous avez compris le chausse trappe, vous saurez pourquoi il faut écrire  $M = \mathbb{C}[X]p$  et non utiliser l'écriture plus simple «  $M = (p)$  ».

(4)  **$AC[X] = A$**

L'inclusion  $A \subset AC[X]$  est due au fait que  $1 \in \mathbb{C}[X]$ , et l'autre inclusion est le fait que  $\mathbb{C}[X] \subset A$  alors que  $A$  est un anneau.

(5) **Suite des opérations**

Nous avons :

$$J = AS \subset AC[X]p. \quad (26.258)$$

Voilà une inclusion de montrée. Reste à faire l'autre.

Vu que  $p \in J$  nous avons aussi  $Ap \subset J$ . Et donc

$$AC[X]p = Ap \subset J. \quad (26.259)$$

Avec ces deux inclusions,  $J = AC[X]p = Ap$ . Donc  $J$  est engendré par un seul élément et est principal.

□

41. Être holomorphe sur un compact signifie qu'il existe une extension holomorphe à un ouvert contenant le compact.

42. Définition 1.340

### 26.2.13 Prolongement de fonctions holomorphes

PropDRnYkKP

#### Proposition 26.69.

Soient  $\Omega$  un ouvert de  $\mathbb{C}$  et  $f: \Omega \rightarrow \mathbb{C}$  une fonction holomorphe sur  $\Omega \setminus \{a\}$  ( $a \in \Omega$ ). Nous supposons qu'il existe  $r > 0$  tel que  $f$  est bornée sur  $B(a, r) \cap \Omega$ . Alors  $f$  se prolonge en une fonction holomorphe sur  $\Omega$ .

Le théorème de prolongement de Riemann 26.91 donnera plus d'informations.

*Démonstration.* Nous définissons la fonction  $g: \Omega \rightarrow \mathbb{C}$  par

$$g(z) = \begin{cases} (z-a)f(z) & \text{si } z \neq a \\ 0 & \text{si } z = a. \end{cases} \quad (26.260)$$

Sur  $\Omega \setminus \{a\}$ , la fonction  $g$  est holomorphe (produit de fonctions holomorphes), et elle est continue en  $a$ . Par conséquent elle est holomorphe sur  $\Omega$ . Nous la développons en série entière sur une boule  $B(a, r)$  :

$$g(z) = \sum_{n=0}^{\infty} c_n (z-a)^n. \quad (26.261)$$

Nous avons  $g(a) = c_0 = 0$ . Nous posons

$$\varphi(z) = \sum_{n=0}^{\infty} c_{n+1} (z-a)^n. \quad (26.262)$$

Si  $z \neq a$ , alors  $\varphi(z) = f(a)$  parce que  $\varphi(z) = g(z)/(z-a)$ . Mais  $\varphi$  est continue en  $a$ , et donc holomorphe en  $a$ .

La fonction  $\varphi$  est par conséquent un prolongement holomorphe de  $f$  en  $a$ .  $\square$

### 26.2.14 Théorème de Runge

Le théorème que nous allons prouver n'est en réalité qu'une partie de ce qui est usuellement appelé le théorème de Runge.

ThoMvMCci

#### Théorème 26.70 (Théorème de Runge).

Soit  $K$ , un compact de  $\mathbb{C}$  tel que  $\mathbb{C}K$  soit connexe. Si  $a \in \mathbb{C}K$  alors la fonction

$$\varphi_a(z) = \frac{1}{z-a} \quad (26.263)$$

est limite uniforme de polynômes sur  $K$ .

*Démonstration.* Nous considérons  $P(K)$ , l'adhérence des polynômes sur  $K$  pour la norme uniforme (sur  $K$ ). Nous devons montrer que pour tout  $a \in \mathbb{C}K$ , la fonction  $\varphi_a$  est dans  $P(K)$ . Pour cela nous considérons l'ensemble

$$A = \{a \in \mathbb{C}K \text{ tel que } \varphi_a \in P(K)\} \quad (26.264)$$

et nous allons montrer qu'il est à la fois non vide, ouvert et fermé dans le connexe  $\mathbb{C}K$ .

Je répète : nous allons prouver l'ouverture et la fermeture *pour la topologie de  $\mathbb{C}K$* . Nous n'allons pas prouver que  $A$  est un ouvert de  $\mathbb{C}$ . Ce qui sera par conséquent prouvé est que  $A = \mathbb{C}K$ .

(1) **Non vide** Soit  $R = \sup_{z \in K} |z|$  et  $a \in \mathbb{C}K$  tel que  $|a| > R$ . Nous avons

$$\varphi_a(z) = \frac{1}{a \frac{z}{a} - 1} = -\frac{1}{a} \frac{1}{1 - \frac{z}{a}} = -\frac{1}{a} \sum_{k=0}^{\infty} \left(\frac{z}{a}\right)^k = -\sum_{k=0}^{\infty} \frac{z^k}{a^{k+1}}. \quad (26.265)$$

Ici la convergence de la série et sa limite sont assurées par le fait que  $|z/a| < 1$  par choix de  $R$  et  $a$ . La suite de polynômes

$$P_n(z) = \sum_{k=0}^n \frac{z^k}{a^{k+1}} \quad (26.266)$$

converge uniformément sur  $B(0, R)$  et en particulier sur  $K$ . Donc  $P_n \rightarrow \varphi_a$ .

(2) **Fermé**

Nous allons montrer que la fermeture de  $A$  (dans  $\mathbb{C}K$ ) est incluse dans  $A$ , et donc qu'elle est égale à  $A$  et donc que  $A$  est fermé. Par le lemme 7.34, la fermeture de  $A$  dans  $\mathbb{C}K$  est l'ensemble  $\bar{A} \cap \mathbb{C}K$  où  $\bar{A}$  est la fermeture de  $A$  au sens usuel.

Bref, soit  $a \in \bar{A} \cap \mathbb{C}K$ , et montrons que  $\varphi_a \in \overline{P(K)}$ . Vu que  $P(K)$  est déjà une fermeture, nous aurons en fait  $\varphi_a \in P(K)$  et donc  $a \in A$ , ce qui signifierait que  $\bar{A} \cap \mathbb{C}A = A$  et donc que  $A$  est fermé.

Au travail.

Soit  $(a_n) \in A$  une suite convergente vers  $a$ . Soit aussi  $d = d(a, K)$ ; on a  $d > 0$  parce que  $K$  est compact et  $a$  est hors de  $K$  alors le complémentaire de  $K$  est ouvert. Nous choisissons en plus la suite  $a_n$  pour avoir  $|a_n - a| < \frac{d}{2}$ ; au pire on prend la queue de suite. Soit  $z \in K$ ; nous avons

$$|\varphi_{a_n}(z) - \varphi_a(z)| = \left| \frac{1}{z - a_n} - \frac{1}{z - a} \right| = \left| \frac{a_n - a}{(z - a_n)(z - a)} \right|. \quad (26.267)$$

Vu que  $a_n \in B(a, \frac{d}{2})$  et que  $z \in K$  et  $d = d(a, K)$  nous avons  $|a_n - z| \geq \frac{d}{2}$ ; et aussi  $|a - z| \geq \frac{d}{2}$ . Nous pouvons donc majorer (26.267) par

$$|\varphi_{a_n}(z) - \varphi_a(z)| \leq 2 \frac{|a_n - a|}{d^2}. \quad (26.268)$$

Donc nous avons

$$\|\varphi_a - \varphi_{a_n}\|_K \leq 2 \frac{|a_n - a|}{d^2} \rightarrow 0 \quad (26.269)$$

où la norme  $\|\cdot\|_K$  est la norme supremum sur  $K$ . Donc  $a \in \overline{P(K)} = P(K)$  et  $A$  est fermé.

(3) **Ouvert** Vu que  $K$  est compact, il est fermé et donc  $\mathbb{C}K$  est ouvert. Par conséquent, ainsi que précisé dans l'exemple 7.41, les ouverts de  $\mathbb{C}K$  sont les ouverts de  $\mathbb{C}$  contenus dans  $\mathbb{C}K$ . Afin de prouver que  $A$  est ouvert, nous prenons  $a \in A$  et nous cherchons une boule (au sens de  $\mathbb{C}$ ) autour de  $a$  qui serait incluse dans  $A$ .

Soit donc  $h \in \mathbb{C}$  « petit » dans un sens que nous allons préciser plus tard. Encore une fois nous posons  $d = d(a, K)$ . Nous avons

$$\varphi_{a+h}(z) = \frac{1}{z - a - h} = \frac{1}{z - a} \frac{1}{1 - \frac{h}{z - a}} = \sum_{k=0}^{\infty} \frac{h^k}{(z - a)^{k+1}}. \quad (26.270)$$

Déjà ici nous demandons  $h < \sup_{z \in K} |z - a|$ . Puisque  $|z - a| > d$ , nous avons alors

$$|\varphi_{a+h}(z)| \leq \sum_{k=0}^{\infty} \frac{h^k}{d^{k+1}} < \infty. \quad (26.271)$$

Cela pour dire que la somme à droite de (26.270) converge bien pourvu que  $h$  soit bien petit. Nous pouvons donc poursuivre :

$$\varphi_{a+h}(z) = \sum_{k=0}^{\infty} \frac{h^k}{(z - a)^{k+1}} = \sum_{k=0}^{\infty} h^k \varphi_a(z)^{k+1}. \quad (26.272)$$

Nous montrons maintenant que la convergence de la somme (26.272) est en réalité uniforme en  $z$ . En effet

$$\left| \varphi_{a+h}(z) - \sum_{k=0}^N h^k \varphi_a(z)^{k+1} \right| = \left| \sum_{k=N+1}^{\infty} h^k \varphi_a(z)^{k+1} \right| \quad (26.273a)$$

$$\leq \sum_{k=N+1}^{\infty} |h|^k |\varphi_a(z)|^{k+1}. \quad (26.273b)$$

Étant donné que  $\varphi_a$  est continue sur le compact  $K$ , elle y est majorée en module ; on peut même être plus précis :

$$|\varphi_a(z)| = \frac{1}{|z-a|} \leq \frac{1}{d}. \quad (26.274)$$

Nous pouvons donc écrire

$$|\varphi_{a+h}(z) - \sum_{k=0}^N h^k \varphi_a(z)^{k+1}| \leq \frac{1}{d} \sum_{k=N+1}^{\infty} \left| \frac{h}{d} \right|^k. \quad (26.275)$$

Étant donné que la somme  $\sum_{k=0}^{\infty} |h/d|^k$  converge, la limite  $N \rightarrow \infty$  est nulle et nous avons

$$\lim_{N \rightarrow \infty} \|\varphi_{a+h} - \sum_{k=0}^N h^k \varphi_a^{k+1}\|_K = 0. \quad (26.276)$$

Pour avoir  $\varphi_{a+h} \in P(K)$ , il faut encore savoir si les fonctions  $\varphi_a^k$  sont dans  $P(K)$  pour tout  $k$ . Dans ce cas pour chaque  $N$  la somme sera encore dans  $P(K)$  et  $\varphi_{a+h}$  sera limite uniforme d'éléments de  $P(K)$ .

Par hypothèse,  $\varphi_a \in P(K)$  ; soit  $P_n$  une suite de polynômes qui converge uniformément vers  $\varphi_a$ . Nous allons montrer qu'alors la suite de polynômes  $P_n^k$  converge uniformément vers  $\varphi_a^k$ . Soit  $n$  tel que  $\|P_n - \varphi_a\|_K \leq \epsilon$  et utilisons le produit remarquable

$$a^k - b^k = (a-b) \sum_{i=0}^{k-1} a^i b^{k-1-i} \quad (26.277)$$

pour obtenir

$$|P_n(z)^k - \varphi_a(z)^k| \leq |P_n(z) - \varphi_a(z)| \sum_{i=0}^{k-1} |P_n(z)^i \varphi_a(z)^{k-1-i}|. \quad (26.278)$$

Vu que  $P_n$  et  $\varphi_a$  sont continues sur le compact  $K$ , on peut majorer la somme par une constante  $M$ , et il restera

$$|P_n(z)^k - \varphi_a(z)^k| \leq M |P_n(z) - \varphi_a(z)|, \quad (26.279)$$

ou encore

$$\|P_n^k - \varphi_a^k\| \leq M\epsilon. \quad (26.280)$$

Cela prouve que  $\varphi_a^k \in P(K)$  et donc que  $\varphi_{a+h}$  est limite uniforme (sur  $K$ ) d'éléments de  $P(K)$  et donc fait partie de  $P(K)$  lui aussi.

Ceci achève de prouver que  $A$  est ouvert dans  $\mathbb{C}K$ .

#### (4) Conclusion

L'ensemble  $A$  est non vide, ouvert et fermé dans  $\mathbb{C}K$ , donc il est égal à  $\mathbb{C}K$ . Le théorème est ainsi démontré. □

## 26.3 Intégrales de fonctions holomorphes

Nous commençons par le lemme technique.

**Lemme 26.71** ([597]).

LemNAnweA

Soit  $f$  une fonction holomorphe sur  $B(z_0, r_0)$ . Pour tout  $z \in B(z_0, r)$  (avec  $r < r_0$ ) nous avons

$$|f'(z)| \leq \frac{r}{(r - |z - z_0|)^2} \max \{f(z_0 + re^{i\theta})\}_{\theta \in \mathbb{R}}. \quad (26.281)$$

*Démonstration.* Par translation nous pouvons supposer que  $z_0 = 0$ . Étant donné que  $f$  est holomorphe, elle admet un développement en séries entières

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad (26.282)$$

et nous notons  $M = \max\{|f(z)| \text{ tel que } z \in \overline{B(0, r)}\}$ . Nous avons  $r^n |a_n| \leq M$ . Par conséquent

$$|f'(z)| = \left| \sum_{n=1}^{\infty} n a_n z^{n-1} \right| \quad (26.283a)$$

$$\leq \frac{1}{r} \sum_{n=1}^{\infty} r^n |a_n| n \left( \frac{|z|}{r} \right)^{n-1} \quad (26.283b)$$

$$< \frac{M}{r} \sum_{n=1}^{\infty} n \left( \frac{|z|}{r} \right)^{n-1}. \quad (26.283c)$$

À ce point nous devons utiliser la série de l'exemple 15.159. Nous avons alors

$$|f'(z)| \leq \frac{M}{r} \frac{1}{\left(1 - \frac{|z|}{r}\right)^2} = \frac{Mr}{(r - |z|)^2}. \quad (26.284)$$

□

### 26.3.1 Holomorphie sous l'intégrale

#### 26.72.

Notez une grande différence entre les théorèmes 26.73 et 17.19 : la condition (17.76) demande de contrôler l'intégrabilité de la dérivée de  $f$ , alors que la condition (26.286) demande de contrôler l'intégrabilité de  $f$  elle-même. Oh oui, on voudrait faire de l'analyse, mais ces fonctions holomorphes ...tellement déloyal!!

ThopCLOVN

**Théorème 26.73** (Holomorphie sous l'intégrale[597]).

Soit un espace mesuré  $(\Omega, \mu)$ , un ouvert  $A$  dans  $\mathbb{C}$  et une fonction  $f: A \times \Omega \rightarrow \mathbb{C}$ . Nous voulons étudier la fonction

$$F(z) = \int_{\Omega} f(z, \omega) d\mu(\omega) \quad (26.285)$$

pour tout  $z \in A$ . Nous supposons que

- (1) la fonction  $f(\cdot, \omega)$  est holomorphe sur  $A$  pour chaque  $\omega$ .
- (2) La fonction  $f(z, \cdot)$  est mesurable sur  $(\Omega, \mu)$ .
- (3) Pour tout compact  $K \subset A$ , il existe une fonction  $g_K: \Omega \rightarrow \mathbb{R}$  telle que  $|f(z, \omega)| \leq g_K(\omega)$  et telle que

$$\int_{\Omega} g_K(\omega) d\mu(\omega) < \infty \quad (26.286)$$

EQooYILMooAIBQof

existe.

Alors la fonction  $F$  est holomorphe et

$$F'(z) = \int_{\Omega} \frac{\partial f}{\partial z}(z, \omega) d\mu(\omega). \quad (26.287)$$

*Démonstration.* Soient  $z_0 \in A$  et  $r > 0$  tels que  $K = \overline{B(z_0, r)} \subset A$ . Pour chaque  $\omega \in \Omega$  nous considérons la fonction

$$f_{\omega}: \overline{B(z_0, r)} \rightarrow \mathbb{C} \\ z \mapsto f(z, \omega). \quad (26.288)$$

Étant donné que  $\overline{B(z_0, r)}$  est compacte, la fonction  $|f_\omega|$  est majorée par un nombre que nous notons  $f_K(\omega)$  qui est indépendant de  $z$  (pour autant que  $z \in K$ ). Nous désignons par  $S(z_0, r)$  la frontière de la boule  $B(z_0, r)$ . Étant donné que la majoration est valable sur  $\overline{B(z_0, r)}$ , nous avons en particulier

$$|f_\omega(z)| \leq f_K(\omega) \quad (26.289)$$

pour tout  $z \in S$ . En utilisant la lemme 26.71 nous avons

$$|f'_\omega(z)| \leq \frac{r}{(r - |z - z_0|)^2} \max\{f(z_0 + re^{i\theta})\}_{\theta \in \mathbb{R}} \quad (26.290a)$$

$$\leq \frac{rf_K(\omega)}{(r - |z - z_0|)^2}. \quad (26.290b)$$

Cette majoration est valable pour tout  $z \in B(z_0, r)$ . Si nous supposons de plus que  $z \in B(z_0, r/2)$  nous avons

$$|f'(z)| \leq \frac{rf_K(\omega)}{(r - \frac{r}{2})^2} = \frac{4}{r} f_K(\omega). \quad (26.291)$$

Étant donné que la boule  $B(z_0, r/2)$  est convexe, la fonction  $f_\omega$  est Lipschitz et pour tout  $h \in \mathbb{C}$  tel que  $|h| < r/2$  nous avons

$$\left| \frac{f_\omega(z_0 + h) - f_\omega(z_0)}{h} \right| \leq \frac{4f_K(\omega)}{r}. \quad (26.292)$$

Soit maintenant une suite  $(h_n)$  qui converge vers 0 dans  $\mathbb{C}$ . Nous considérons la suite de fonctions correspondantes

$$g_n(\omega) = \frac{f(z_0 + h_n, \omega) - f(z_0, \omega)}{h_n}. \quad (26.293)$$

Cette suite de fonctions vérifie la convergence ponctuelle

$$g_n(\omega) \rightarrow \frac{\partial f}{\partial z}(z_0, \omega). \quad (26.294)$$

De plus  $g_n$  est une fonction (de  $\omega$ ) dominée par  $\frac{4f_K}{r}$  qui est intégrable. Par conséquent le théorème de la convergence dominée<sup>43</sup> nous indique que

$$\int_{\Omega} g_n(\omega) d\mu(\omega) \rightarrow \int_{\Omega} \frac{\partial f}{\partial z}(z_0, \omega) d\mu(\omega), \quad (26.295)$$

tandis que

$$F'(z) = \lim_{n \rightarrow \infty} \frac{F(z_0 + h_n) - F(z_0)}{h_n} = \lim_{n \rightarrow \infty} \int_{\Omega} g_n(\omega) d\mu(\omega). \quad (26.296)$$

□

#### 26.74.

Dans le corolaire suivant, l'intégrale sur  $\partial B(z, r)$  est une intégrale sur un chemin.

CorNxTjEj

#### Corolaire 26.75.

Si  $f$  est une fonction holomorphe sur un ouvert  $\Omega$  contenant la fermeture de la boule  $B(z_0, r)$ , alors pour tout  $z$  dans  $B(z_0, \rho)$  ( $\rho < r$ ) les dérivées de  $f$  s'expriment par la formule suivante :

$$f^{(k)}(z) = \frac{k!}{2\pi i} \int_{\partial B(z_0, r)} \frac{f(\xi)}{(\xi - z)^{k+1}} d\xi. \quad (26.297)$$

EQooBPIQooNhOTtE

*Démonstration.* Nous faisons par récurrence.

43. Lebesgue, théorème 14.207.



- (1) **Pour la dérivée première** Soit  $\rho < r$ . Nous utilisons le théorème 26.60 pour écrire

$$f(z) = \frac{1}{2\pi i} \int_{\partial B(z_0, r)} \frac{f(\xi)}{\xi - z} d\xi. \quad (26.298)$$

Ensuite nous dérivons sous l'intégrale en appliquant le théorème 26.73 à la fonction

$$\begin{aligned} g: B(z_0, \rho) \times \partial B(z_0, r) &\rightarrow \mathbb{C} \\ (z, \xi) &\mapsto \frac{f(\xi)}{\xi - z}. \end{aligned} \quad (26.299)$$

Étant donné que  $f$  est holomorphe, la fonction  $g$  est continue et donc bornée sur tout compact  $K \subset A$  par une constante  $M$  (qui dépend du compact choisi). D'autre part, nous avons toujours  $|\xi - z| > r - \rho$  et donc

$$|g(z, \xi)| \leq \frac{M}{r - \rho}. \quad (26.300)$$

La fonction constante  $g_K = \frac{M}{r - \rho}$  est évidemment intégrable. Le théorème conclut que  $f$  est holomorphe (cela, nous le savions déjà<sup>44</sup>), et

$$f'(z) = \frac{1}{2i\pi} \int_{\partial B} \frac{f(\xi)}{(\xi - z)^2} d\xi. \quad (26.301)$$

- (2) **Les dérivées suivantes** Pour la récurrence[701] nous supposons que

$$f^{(k)}(z) = \frac{k!}{2i\pi} \int_{\partial B(z_0, r)} \frac{f(\xi)}{(\xi - z)^{k+1}} d\xi, \quad (26.302)$$

et nous tentons de calculer  $f^{(k+1)}(z)$ . Pour cela nous paramétrisons l'intégrale de façon très usuelle :

$$f^{(k)}(z) = \frac{k!}{2\pi i} \int_0^{2\pi} \frac{f(re^{it})}{(re^{it} - z)^{k+1}} ire^{it} dt. \quad (26.303)$$

Nous permettons de permuter la  $\mathbb{C}$ -dérivation (par rapport à  $z$ ) et l'intégrale en vertu de la proposition 26.64 appliquée à la fonction

$$g(t, z) = \frac{f(re^{it})}{(re^{it} - z)^{k+1}} ire^{it}. \quad (26.304)$$

Cela donne

$$f^{(k+1)}(z) = \frac{k!}{2i\pi} \int_0^{2\pi} f(re^{it}) ire^{it} \frac{k+1}{(re^{it} - z)^{k+1}} dt = \frac{(k+1)!}{2i\pi} \int_{\partial B} \frac{f(\xi)}{(\xi - z)^{k+2}} d\xi. \quad (26.305)$$

□

THOoSULFooHTLRPE

### **Théorème 26.76.**

Si  $f$  est une fonction holomorphe sur le disque ouvert  $B(z_0, R)$  alors

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n \quad (26.306)$$

et cette série converge uniformément sur tout compact.

<sup>44</sup>. Et cela fournit une preuve alternative à la réciproque du théorème de Cauchy : une fonction continue qui vérifie la formule de Cauchy est holomorphe.

*Démonstration.* Sans perte de généralité nous supposons que  $z_0 = 0$ . La formule de Cauchy (théorème 26.60) fournit, pour  $z \in B(0, R)$ ,

$$f(z) = \frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{\xi - z} d\xi = \frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{1 - (z/\xi)} \frac{d\xi}{\xi}. \quad (26.307)$$

En particulier notons que  $z \in B(0, R)$  alors que  $\xi$  est sur le bord de cette boule ouverte. Donc  $|\xi| > |z|$  pour tous les  $\xi$  et  $z$  qui interviennent. Nous utilisons la série géométrique

$$\frac{1}{1 - (z/\xi)} = \sum_{n=0}^{\infty} \left(\frac{z}{\xi}\right)^n. \quad (26.308)$$

- (1) Permuter une intégrale et une somme En utilisant la mesure de comptage<sup>45</sup> sur  $\mathbb{N}$  (qui est  $\sigma$ -finie), nous pouvons écrire

$$\int_{\partial B} \sum_{n=0}^{\infty} \frac{z^n f(\xi)}{\xi^{n+1}} d\xi = \int_{\partial B} \left( \int_{\mathbb{N}} g(\xi, n) dm(n) \right) d\xi \quad (26.309)$$

où

$$\begin{aligned} g: \partial B \times \mathbb{N} &\rightarrow \mathbb{C} \\ (\xi, n) &\mapsto \frac{z^n f(\xi)}{\xi^{n+1}}. \end{aligned} \quad (26.310)$$

Nous allons permuter les intégrales en utilisant le théorème de Fubini, selon la procédure décrite en 14.306. Nous commençons par l'intégrale sur  $\mathbb{N}$  :

$$\int_{\mathbb{N}} |g(n, \xi)| = \left| \frac{f(\xi)}{\xi} \right| \sum_{n \in \mathbb{N}} \left| \frac{z}{\xi} \right|^n = \frac{1}{R} |f(\xi)| \frac{1}{1 - |z|/R}. \quad (26.311)$$

Ici nous avons utilisé  $|\xi| = R$ . Notons que  $z$  est fixé depuis longtemps à l'intérieur de la boule de rayon  $R$  de telle sorte que  $|z/\xi|$  est une constante strictement inférieure à 1.

L'intégrale sur  $\xi \in \partial B$  n'a pas à être effectuée explicitement : nous nous contentons de prouver qu'elle est finie. La fonction  $f$  est continue sur le compact  $\partial B$ . Cela parce que  $B$  est une boule fermée dans l'ouvert  $\Omega$  sur lequel  $f$  est continue. Au final l'expression à droite de (26.311) est bornée sur le compact  $\partial B$  et son intégrale donne un nombre fini.

Tout ceci pour invoquer le corolaire 14.304 qui nous indique que  $g \in L^1(\mathbb{N} \times \partial B)$ .

Une fois  $g$  intégrable sur l'espace produit  $\mathbb{N} \times \partial B$ , nous pouvons utiliser Fubini 14.305 pour permuter les intégrales.

Une fois la somme et l'intégrale permutées, nous avons

$$f(z) = \frac{1}{2\pi i} \sum_{n=0}^{\infty} \int_{\partial B} \frac{z^n f(\xi)}{\xi^{n+1}} = \sum_{n=0}^{\infty} \left( \frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{\xi^{n+1}} \right) z^n. \quad (26.312)$$

Nous devons maintenant montrer que ce qui se trouve dans la grande parenthèse vaut  $f^{(n)}(0)/n!$ . Cela est immédiat en comparant avec la formule (26.297).

□

ThoRckxes

**Proposition 26.77** (Morera [702]).

Soit  $\Omega$  ouvert dans  $\mathbb{C}$  et  $f$  continue. Si

$$\int_{\partial T} f = 0 \quad (26.313)$$

pour tout triangle (plein)  $T$  contenu dans  $\Omega$ , alors  $f$  est holomorphe sur  $\Omega$ .

45. Définition 14.264.

*Démonstration.* Il est suffisant de prouver que  $f$  est holomorphe sur toute boule ouverte  $B(a, r)$  incluse dans  $\Omega$ . Nous posons, pour tout  $z \in B(a, r)$ ,

$$F(z) = \int_{[a, z]} f, \quad (26.314)$$

et nous considérons le chemin triangulaire  $a \rightarrow z \rightarrow z + h \rightarrow a$  où  $h \in \mathbb{C}$  est choisi assez petit pour que  $z + h \in B(a, r)$ . L'intégrale sur le triangle étant nulle, nous avons

$$0 = \int_{a \rightarrow z} f + \int_{z \rightarrow z+h} f + \int_{z+h \rightarrow a} f, \quad (26.315)$$

c'est-à-dire

$$F(z + h) - F(z) = \int_{z \rightarrow z+h} f. \quad (26.316)$$

En paramétrant le chemin par  $z + th$  avec  $t \in [0, 1]$ , et en tenant compte de la remarque 20.44,

$$F'(z) = \lim_{h \rightarrow 0} \frac{F(z + h) - F(z)}{h} \quad (26.317a)$$

$$= \lim_{h \rightarrow 0} \frac{1}{h} \int_0^1 f(z + th) h dt, \quad (26.317b)$$

ce qui prouve que  $F$  est dérivable et  $F' = f$ . Par définition (12.331),  $F$  est holomorphe, et donc  $C^\infty$  par le théorème 26.61. Du coup  $f$  est également  $C^\infty$  et donc en particulier holomorphe.  $\square$

### 26.3.2 Mesure de Radon

Soit un compact  $K$  de  $\mathbb{C}$ . Le lemme 12.384 nous dit que  $(C(K), \|\cdot\|_\infty)$  est complet. Mais cela n'est pas vraiment utile pour l'instant. Je dis ça juste pour le garder dans un coin de la tête.

DEFooTLRKooGbKkAj

#### Définition 26.78.

Une **mesure de Radon** sur un compact  $K$  de  $\mathbb{C}$  est une forme linéaire continue sur  $(C(K), \|\cdot\|_\infty)$ . Si  $\mu$  est une mesure de Radon, on définit la **transformée de Cauchy** de  $\mu$  par

$$\begin{aligned} \hat{\mu}: \mathbb{C} \setminus K &\rightarrow \mathbb{C} \\ z &\mapsto -\frac{1}{\pi} \mu(f_z) \end{aligned} \quad (26.318)$$

où, pour tout  $z \in \mathbb{C} \setminus K$  nous avons défini la fonction  $f_z$  par

$$\begin{aligned} f_z: K &\rightarrow \mathbb{C} \\ \xi &\mapsto \frac{1}{\xi - z}. \end{aligned} \quad (26.319)$$

EQooNHTLooWkIofe

#### 26.79.

Les fonctions  $f_z$  sont définies pour  $z \in \mathbb{C} \setminus K$ , et elles sont définies sur  $K$ . Donc il n'y a pas de danger que  $\xi = z$  dans le dénominateur de (26.319). De plus si  $f_z$  est définie (c'est-à-dire si  $z \in \mathbb{C} \setminus K$ ), alors  $f_{z+h}$  est également définie tant que  $h \in \mathbb{C}$  n'est pas trop grand parce que  $\mathbb{C} \setminus K$  est un ouvert de  $\mathbb{C}$ .

ThoJVNTzn

#### Théorème 26.80.

Si  $\mu$  est une mesure de Radon<sup>46</sup> sur  $K$  alors  $\hat{\mu}$  est infiniment  $\mathbb{C}$ -dérivable<sup>47</sup> sur  $\Omega = \mathbb{C} \setminus K$  et nous avons

$$\hat{\mu}^{(n)}(z) = -\frac{n!}{\pi} \mu \left( \frac{1}{(\xi - z)^{n+1}} \right). \quad (26.320)$$

46. Définition 26.78.

47. Définition 12.330.

*Démonstration.* Calculons la dérivée complexe de  $\hat{\mu}$ . Si la limite existe, nous avons

$$\hat{\mu}'(z) = \lim_{\substack{h \rightarrow 0 \\ h \in \mathbb{C}}} \frac{\hat{\mu}(z+h) - \hat{\mu}(z)}{h}. \quad (26.321)$$

Nous allons nous concentrer sur ce qui est dans la limite pour montrer qu'elle existe, et pour calculer la valeur. Nous avons

$$\frac{\hat{\mu}(z+h) - \hat{\mu}(z)}{h} = -\frac{1}{\pi} \frac{\mu(f_{z+h}) - \mu(f_z)}{h} \quad (26.322a)$$

$$= -\frac{1}{\pi} \mu \left( \frac{f_{z+h} - f_z}{h} \right) \quad \mu \text{ est linéaire} \quad (26.322b)$$

$$= -\frac{1}{\pi} \mu \left( \frac{1}{(\xi - z - h)(\xi - z)} \right) \quad (26.322c)$$

où la dernière ligne est un abus de notations pour dire que  $\mu$  est appliquée à la fonction  $\xi \mapsto 1/(\xi - z - h)(\xi - z)$ .

Histoire d'éviter de regrettables abus de notations, nous introduisons la fonction

$$\begin{aligned} \phi_{z,h}: K &\rightarrow \mathbb{C} \\ \xi &\mapsto \frac{1}{(\xi - z - h)(\xi - z)}, \end{aligned} \quad (26.323)$$

ainsi que la famille

$$\begin{aligned} f_{z,n}: K &\rightarrow \mathbb{C} \\ \xi &\mapsto \frac{1}{(\xi - z)^n}. \end{aligned} \quad (26.324)$$

Ce que nous avons montré est que

$$\frac{\hat{\mu}(z+h) - \hat{\mu}(z)}{h} = -\frac{1}{\pi} \mu(\phi_{z,h}). \quad (26.325)$$

De plus nous avons la convergence

$$\phi_{z,h} \xrightarrow{(C(K), \|\cdot\|_\infty)} f_{z,2} \quad (26.326)$$

lorsque  $h \rightarrow 0$ . Vu que  $\mu$  est continue, nous pouvons permuter la limite avec  $\mu$  et avoir :

$$\hat{\mu}'(z) = \lim_{h \rightarrow 0} \mu(\phi_{z,h}) = -\frac{1}{\pi} \mu(f_{z,2}). \quad (26.327)$$

Cela prouve le résultat pour  $n = 1$ .

Pour le reste, vous faites un récurrence <sup>48</sup>. □

Cette théorie permet de fournir une démonstration plus technologique du corolaire 26.75.

### Lemme 26.81.

Si  $f$  est holomorphe sur  $\Omega$  et si  $B$  est une boule fermée dans  $\Omega$  alors pour tout  $z \in \text{Int}(B)$  nous avons

$$f^{(k)}(z) = \frac{k!}{2i\pi} \int_{\partial B} \frac{f(\xi)}{(\xi - z)^{k+1}} d\xi. \quad (26.328)$$

*Démonstration.* Appliquer le théorème 26.80 à la mesure de Radon

$$\mu(\phi) = \int_{\partial B} \phi(\xi) d\xi. \quad (26.329)$$

□

---

48. Je n'ai pas vérifié. Faites-le et écrivez-moi pour me dire si c'est bon.

Tout ce petit monde à propos de la mesure de Radon permet également de redémontrer que

$$\left( \frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{\xi^{n+1}} \right) = f^{(n)}(0)/n!, \quad (26.330)$$

comme nous l'avons déjà fait autour de l'équation (26.312). Nous utilisons le théorème de Radon 26.80 à la mesure

$$\mu(\phi) = \int_{\partial B} \phi(\xi) d\xi. \quad (26.331)$$

La transformée de Cauchy est

$$\hat{\mu}(z) = -\frac{1}{\pi} \mu \left( \frac{1}{\xi - z} \right) = -\frac{1}{\pi} \int_{\partial B} \frac{1}{\xi - z} d\xi, \quad (26.332)$$

et le théorème assure que

$$\hat{\mu}^{(n)}(z) = -\frac{n!}{\pi} \mu \left( \frac{1}{(\xi - z)^{n+1}} \right) = -\frac{n!}{\pi} \int_{\partial B} \frac{1}{(\xi - z)^{n+1}} d\xi. \quad (26.333)$$

En comparant la formule (26.332) avec la formule de Cauchy nous voyons que  $\hat{\mu}(z) = -2if(z)$ . Par conséquent

$$f^{(n)}(z) = -\frac{1}{2i} \hat{\mu}^{(n)}(z) = \frac{n!}{2\pi i} \int_{\partial B} \frac{1}{(\xi - z)^{n+1}} d\xi, \quad (26.334)$$

et

$$f^{(n)}(0) = \frac{n!}{2\pi i} \int_{\partial B} \frac{1}{\xi^{n+1}} d\xi. \quad (26.335)$$

## 26.4 Conditions équivalentes à l'holomorphic

Nous nous proposons de lister les conditions que nous avons vues être équivalentes à l'holomorphic.

### Théorème 26.82.

Soient  $\Omega$  un ouvert de  $\mathbb{C}$  et  $f: \Omega \rightarrow \mathbb{C}$  une fonction continue. Les conditions suivantes sont équivalentes.

- (1)  $f$  est holomorphic.
- (2) Pour tout triangle (plein)  $T$  contenu dans  $\Omega$ ,  $\int_T f = 0$ .
- (3)  $f$  est  $\mathbb{C}$ -dérivable.
- (4)  $f$  est  $C^\infty$
- (5)  $\partial_{\bar{z}} f = 0$ ; ce sont les équations de Cauchy-Riemann.
- (6) La 1-forme différentielle  $f(z)dz$  est localement exacte.
- (7) Pour toute boule  $B(a, r)$  contenue dans  $\Omega$  nous avons

$$f(a) = \frac{1}{2\pi i} \int_{\partial B(a, r)} \frac{f(z)}{z - a} dz. \quad (26.336)$$

La fonction  $f$  est holomorphic en  $z_0$  si et seulement si il existe un voisinage  $B(z_0, r)$  de  $z_0$  et des nombres  $a_k$  tels que sur la boule,

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \quad (26.337)$$

Dans ce cas,  $f$  est holomorphic sur toute la boule.

*Démonstration.* (1) implique (2) est le lemme de Goursat 26.10. (2) implique (1) est le théorème de Morera 26.77.

(3) est la définition de l'holomorphic, définition 12.331.

(4) implique (1) est un a fortiori sur la définition. (1) implique (4) est contenu dans le théorème de développement en série entière 26.61.

L'équivalence entre (5) et l'holomorphic est le théorème 26.2.

L'équivalence entre (6) et (1) est la proposition 26.65.

L'équivalence entre (1) et (7) est d'une part le théorème 26.61 et d'autre part le corolaire 26.63.

En ce qui concerne la dernière affirmation, si  $f$  est holomorphic en  $z_0$ , alors le théorème 26.61(1) donne la série. Si au contraire nous avons la série, la proposition 26.12 nous donne le résultat grâce au point (3).  $\square$

## 26.5 Limites, séries et dérivée

PROPOOKHBHooAIsbCG

**Proposition 26.83** ([703]).

Soit un ouvert  $\Omega \subset \mathbb{C}$ . Soit une suite  $(f_n)$  de fonctions holomorphes sur  $\Omega$ . Nous supposons que  $f_n$  converge uniformément sur tout compact vers une fonction  $f$ .

Alors

(1)  $f$  est holomorphic.

(2) Pour tout  $k \geq 1$  nous avons  $f_n^{(k)} \rightarrow f^{(k)}$  uniformément sur tout compact.

PROPOOMIDPOoUgSXRp

**Proposition 26.84** ([1]).

Soit une suite  $(f_n: \Omega \rightarrow \mathbb{C})_{n \in \mathbb{N}}$  de fonctions holomorphes. Si la série

$$f(z) = \sum_{n=0}^{\infty} f_n(z) \quad (26.338)$$

converge uniformément, la fonction  $f$  est holomorphic sur  $\omega$ .

De plus

$$f'(z) = \sum_{n=0}^{\infty} f'_n(z). \quad (26.339)$$

## 26.6 Singularités, pôles et méromorphe

DEFooKWDUooVPvtpy

**Définition 26.85.**

Si  $f$  est holomorphic<sup>49</sup> sur un ouvert  $\Omega$ , alors une **singularité** de  $f$  est un point isolé du bord de  $\Omega$ .

(1) La singularité est **effaçable** si la fonction  $f$  s'y prolonge en une fonction holomorphic.

(2) La singularité  $a$  est **isolée** si  $f$  est holomorphic sur  $B(a, r) \setminus \{a\}$ .

DEFooUIJTtooUJFiDG

**Définition 26.86** (pôle d'une fonction[1, 704]).

Soient un ouvert  $\Omega$  de  $\mathbb{C}$  ainsi que  $a \in \Omega$ . La fonction  $f: \Omega \setminus \{a\} \rightarrow \mathbb{C}$  a un **pôle d'ordre**  $n$  en  $a$  si il existe  $r > 0$  et une fonction holomorphic  $g: B(a, r) \rightarrow \mathbb{C}$  telle que

(1)  $g(a) \neq 0$

(2) pour tout  $z \in B(a, r) \setminus \{a\}$  nous avons

$$f(z) = \frac{g(z)}{(z-a)^n}. \quad (26.340)$$

---

49. Définition 12.331.

**Lemme 26.87.**

Soit  $n \in \mathbb{N}$ . Nous notons  $U_n$  l'ensemble des racines  $n^e$  de l'unité<sup>50</sup> dans  $\mathbb{C}$ . La fonction

$$\begin{aligned} f: \mathbb{C} \setminus U_n &\rightarrow \mathbb{C} \\ z &\mapsto \frac{1}{z^n - 1} \end{aligned} \quad (26.341)$$

est holomorphe et possède un pôle d'ordre 1 en chaque point de  $U_n$ .

*Démonstration.* Le fait que  $f$  soit holomorphe est simplement le fait que sur le domaine, le dénominateur est un bête polynôme qui ne s'annule pas.

Nous énumérons  $U_n = \{\xi_i\}_{i=1,\dots,n}$ . Prouvons que  $\xi_k$  est un pôle d'ordre 1 de  $f$ . La première égalité du lemme 19.20 donne  $z^n - 1 = \prod_i (z - \xi_i)$ . D'abord nous considérons  $r > 0$  tel que  $B(\xi_k, r) \cap U_n = \{\xi_k\}$ .

Nous posons

$$\begin{aligned} g_k: B(\xi_k, r) \setminus \{\xi_k\} &\rightarrow \mathbb{C} \\ z &\mapsto \frac{1}{\prod_{i \neq k} (z - \xi_i)}. \end{aligned} \quad (26.342)$$

Cela est bien une fonction holomorphe et nous avons

$$f(z) = \frac{g_k(z)}{(z - \xi_k)}. \quad (26.343)$$

□

**⚠ Avertissement/question à la lectrice !! 26.88**

Je ne suis pas certain que la proposition suivante soit vraie. Écrivez-moi si vous avez un avis.

PROPOoKXXSooGzbmeo

**Proposition 26.89.**

La fonction

$$z \mapsto \begin{cases} \frac{\sin(z)}{z} & \text{si } z \neq 0 \\ 1 & \text{si } z = 0 \end{cases} \quad (26.344)$$

est continue sur un voisinage de zéro dans  $\mathbb{C}$ .

PROPOoQJDCooYwmkRS

**Proposition 26.90.**

Une singularité de  $f$  est un pôle si et seulement si

$$\lim_{z \rightarrow Z} f(z) = \infty. \quad (26.345)$$

Le théorème suivant complète la proposition 26.69.

ThoTLQOEwW

**Théorème 26.91** (Prolongement de Riemann[705]).

Soient un ouvert  $\Omega \subset \mathbb{C}$ , un point  $a \in \Omega$  et une fonction holomorphe  $f: \Omega \setminus \{a\} \rightarrow \mathbb{C}$ . Nous supposons que  $a$  est une singularité<sup>51</sup> de  $f$ . Les points suivants sont équivalents.

ITEMooMLXJooMfuiFN

(1) la singularité  $a$  est effaçable<sup>52</sup> ;

ITEMooBWPEooEltHAa

(2)  $f$  possède un prolongement continu en  $a$  ;

ITEMooEAUOooIWcxHS

(3) il existe un voisinage épointé de  $a$  sur lequel  $f$  est bornée ;

ITEMooETRWooDTTpxs

(4)  $\lim_{z \rightarrow a} (z - a)f(z) = 0$ .

*Démonstration.* En plusieurs implications.

(1) **(1) implique (2)** La fonction  $f$  admet même un prolongement holomorphe.

50. Voir la définition 19.1 et le lemme 19.2.

51. Singularité et singularité effaçable : définition 26.85.

52. Définition 26.85.

- (2) **(2) implique (3)** Soit un prolongement continu  $\tilde{f}: B(a, r) \rightarrow \mathbb{C}$  de  $f$ . La restriction  $\tilde{f}: \overline{B(a, r/2)} \rightarrow \mathbb{C}$  est continue sur un compact et donc bornée<sup>53</sup> tout en étant égale à  $f$  sur  $B(a, r/2) \setminus \{a\}$ .
- (3) **(3) implique (4)** Nous supposons que  $f: B(a, r) \setminus \{a\} \rightarrow \mathbb{C}$  est bornée. Disons  $|f(z)| < A$ . Alors pour tout  $z \in B(a, r) \setminus \{a\}$  nous avons

$$|(z - a)f(z)| \leq A|z - a| \quad (26.346)$$

Or  $\lim_{z \rightarrow a} A(z - a)$  existe et vaut zéro. Donc  $\lim_{z \rightarrow a} |(z - a)f(z)|$  existe et vaut également zéro.

- (4) **(4) implique (1) si  $a = 0$**

Nous commençons par supposer que  $a = 0$ , et nous posons  $D = B(0, r) \setminus \{0\}$ . La fonction  $f: D \rightarrow \mathbb{C}$  est holomorphe et vérifie  $\lim_{z \rightarrow 0} zf(z) = 0$ .

Nous considérons la fonction suivante :

$$g: B(0, r) \rightarrow \mathbb{C} \\ z \mapsto \begin{cases} 0 & \text{si } z = 0 \\ z^2 f(z) & \text{sinon.} \end{cases} \quad (26.347)$$

La fonction  $g$  est holomorphe sur  $D$  parce que  $f$  l'est. Voyons que  $g$  est dérivable en zéro. Pour tout  $z$  sur un voisinage,

$$\frac{g(z) - g(0)}{z} = \frac{z^2 f(z)}{z} = z f(z). \quad (26.348)$$

Or par hypothèse  $\lim_{z \rightarrow 0} zf(z) = 0$  donc  $g'(0) = 0$ , et  $g$  est holomorphe en 0 (c'est la définition 12.331 d'une fonction holomorphe). Bref,  $g$  est holomorphe sur  $B(0, r)$ .

Nous pouvons donc développer  $g$  en série entière<sup>54</sup> dans un voisinage  $B(0, r)$  :

$$g(z) = \sum_{n=0}^{\infty} a_n z^n. \quad (26.349)$$

En utilisant le 26.62,  $a_0 = g(0) = 0$  et  $a_1 = g'(0) = 0$ . Donc en réalité

$$g(z) = \sum_{n=2}^{\infty} a_n z^n. \quad (26.350)$$

Considérons la série entière

$$\sum_{n=0}^{\infty} b_n z^n \quad (26.351)$$

avec  $b_n = a_{n+2}$ . Le lemme 15.17 dit que son rayon de convergence est le même que celui de  $g$ . Donc la fonction

$$h: B(0, r) \rightarrow \mathbb{C} \\ z \mapsto \sum_{n=2}^{\infty} a_n z^{n-2} \quad (26.352)$$

est holomorphe.

Par ailleurs, sur la partie  $D$  (qui ne contient pas  $z = 0$ ) nous pouvons écrire

$$f(z) = \frac{g(z)}{z^2} \quad (26.353)$$

et donc

$$f(z) = \sum_{n=2}^{\infty} a_n z^{n-2}. \quad (26.354)$$

Autrement dit  $f = h$  sur  $D$ , et  $h$  en est un prolongement holomorphe.

53. Théorème 7.196.

54. Théorème 26.61.



(5) **(3) implique (4)**

Nous posons  $g(z) = f(z + a)$ . Nous avons

$$\lim_{z \rightarrow 0} zg(z) = \lim_{z \rightarrow 0} zf(z + a) = \lim_{z \rightarrow -a} (z - a)f(z) = 0. \quad (26.355)$$

Le changement de variable dans la limite est le lemme 7.230. Donc le premier cas s'applique à  $g$  et nous avons un prolongement holomorphe  $\tilde{g}: B(0, r) \rightarrow \mathbb{C}$  de  $g$ . La fonction donnée par  $\tilde{f}(z) = \tilde{g}(z - a)$  prolonge  $f$ .

□

DEFooRMLGooGzRfgn

**Définition 26.92** (Fonction méromorphe[706]).

Soient  $\mathcal{U}$  un ouvert de  $\mathbb{C}$  et  $\{p_i\}$  une suite de points dans  $\mathcal{U}$  sans points d'accumulation (éventuellement il y a un nombre fini de  $p_i$ ). Si la fonction  $f$  est holomorphe sur  $\mathcal{U} \setminus \{p_i\}$  et si chaque  $p_i$  est un point régulier ou un pôle de  $f$ , alors nous disons que  $f$  est **méromorphe** sur  $\mathcal{U}$ .

LEMooCSAFooTYasYM

**Lemme 26.93.**

Soient  $0 < a < b < \infty$ . Nous considérons l'équation différentielle

$$\begin{cases} y'(t) = \frac{n}{t}y(t) \\ y(a) = y_a \end{cases} \quad (26.356a)$$

$$(26.356b)$$

pour la fonction  $y: ]a, b[ \rightarrow \mathbb{R}$ .

L'unique solution est

$$y(t) = \frac{y_a}{a^n} t^n. \quad (26.357)$$

EQooKPYIooMHPJBP

Note :  $a \neq 0$  de toutes façons, donc pas de problèmes.

*Démonstration.* En termes du théorème de Cauchy-Lipschitz 17.39, nous avons  $y'(t) = f(t, y(t))$  avec

$$\begin{aligned} f: ]a, b[ \times \mathbb{R} &\rightarrow \mathbb{R} \\ (t, y) &\mapsto \frac{n}{t}y. \end{aligned} \quad (26.358)$$

Cette fonction  $f$  est Lipschitz et tout ce qu'on veut parce que  $t = 0$  est hors de son domaine ; la régularité de  $f$  peut être étudiée sur le compact  $[a/2, b]$ .

Un calcul direct vérifie que la solution proposée (26.357) est bien une solution. Le théorème de Cauchy-Lipschitz dit qu'elle est unique. □

**26.94.**

Vous voulez savoir comment on trouve la solution  $y(t) = Kt^n$  ? Allez, on vous la fait un peu détendue, sans trop regarder les détails. D'abord,

$$\frac{y'(t)}{y(t)} = \frac{n}{t}. \quad (26.359)$$

Nous intégrons des deux côtés :

$$\int_a^x \frac{y'(t)}{y(t)} dt = n[\ln(t)]_a^x. \quad (26.360)$$

EQooWCQBQooKqifVj

À gauche nous posons  $u(t) = \ln(y(t))$  et nous avons

$$\int_a^x \frac{y'(t)}{y(t)} dt = \int_a^x u'(t) dt = [u(t)]_a^x = \ln(y(x)) - \ln(y(a)). \quad (26.361)$$

Nous mettons ça à la place du membre de gauche de (26.360), nous mettons toutes les constantes dans un  $L$  (en ne nous posant aucune question sur le fait que ce soit positif, nul, que ça peut rentrer dans un logarithme ou non) et :

$$\ln(y(x)) = n \ln(x) + L = \ln(Kx^n). \quad (26.362)$$

où  $K = e^L$ . Et voilà.

LEMooCESMooJRqFvG

**Lemme 26.95** ([1]).

Soit un ouvert  $\Omega$  de  $\mathbb{C}$ . Si  $f: \Omega \rightarrow \mathbb{C}$  est méromorphe sur tout compact de  $\Omega$ , alors  $f$  est méromorphe sur  $\Omega$ .

*Démonstration.* Soit  $z_0 \in \Omega$ . En prenant une suite exhaustive<sup>55</sup> de compacts pour  $\mathbb{C}$  et en prenant les intersections avec  $\Omega$ , il existe un compact  $K$  tel que  $z_0 \in \text{Int}(K) \subset \Omega$ . Vu que  $f$  est méromorphe sur  $K$ , le point  $z_0$  n'est pas un point d'accumulation des pôles de  $f$ . Cela étant valable pour tout  $z_0 \in \Omega$ , nous en déduisons que  $f$  n'a pas de points d'accumulation de pôles. Donc  $f$  est méromorphe.  $\square$

PropPUZTQKI

**Proposition 26.96** ([237]).

Soient  $\Omega$  un ouvert de  $\mathbb{C}$  et  $f_n: \Omega \rightarrow \mathbb{C}$  une suite de fonctions méromorphes telles que pour tout compact  $K$  de  $\Omega$  il existe  $N_K \geq 0$  tel que

- (1)  $f_n$  n'a pas de pôle<sup>56</sup> dans  $K$  dès que  $n \geq N_K$  ;
- (2) la série  $\sum_{n \geq N_K} f_n$  converge uniformément sur  $K$ .

Alors

ITEMooZKACooLKHmaj

- (1) La fonction

$$f(z) = \sum_{n=0}^{\infty} f_n(z) \quad (26.363)$$

est méromorphe<sup>57</sup> sur  $\Omega$  et ses pôles sont dans l'union de ceux des  $f_n$ .

ITEMooRRCWooHXgpoc

- (2) Nous pouvons permuter la somme et la dérivée :

$$f'(z) = \sum_{n=0}^{\infty} f'_n(z) \quad (26.364)$$

pour tout  $z$  en dehors des pôles.

*Démonstration.* En deux parties.

- (1) **Pour (1)** Soit un compact  $K$  dans  $\Omega$ . Les applications  $(f_n)_{n \geq N_K}$  sont holomorphes sur  $K$ . Donc l'application

$$g_K(z) = \sum_{n \geq N_K} f_n(z) \quad (26.365)$$

est holomorphe comme limite uniforme de fonctions holomorphes (proposition 26.84).

La somme finie

$$h_K(z) = \sum_{n=0}^{N_K-1} f_n(z) \quad (26.366)$$

est méromorphe. Donc la somme

$$f(z) = h_K(z) + g_K(z) \quad (26.367)$$

est méromorphe sur  $K$  et ses pôles sont parmi les pôles des  $(f_n)_{n \leq N_K}$ . Elle est donc méromorphe sur  $\Omega$  par le lemme 26.95.

- (2) **Pour (2)** Nous notons  $P$  l'ensemble de tous les pôles des  $f_n$ . C'est un ensemble sans point d'accumulation. Donc  $\Omega \setminus P$  est encore un ouvert. Soit un compact  $K$  dans  $\Omega \setminus P$ . La fonction

$$\begin{aligned} g: K &\rightarrow \mathbb{C} \\ z &\mapsto \sum_{n \geq N_K} f_n(z) \end{aligned} \quad (26.368)$$

55. Lemme 7.358.

56. Définition 26.86.

57. Définition 26.92.

est holomorphe comme précédemment. Par la proposition 26.84, elle vérifie

$$g'(z) = \sum_{n \geq N_K} f'_n(z). \quad (26.369)$$

En tant que somme finie, la fonction  $h(z) = \sum_{n=0}^{N_K-1} f_n$  est également holomorphe sur  $K$ . Au final la série complète  $f(z) = g(z) + h(z)$  est holomorphe sur  $K$  et peut être dérivée terme à terme :

$$f'(z) = g'(z) + h'(z) = \sum_{n=0}^{\infty} f'_n(z). \quad (26.370) \quad \text{EQooFYIBooSzKCVi}$$

Le tout étant valable pour tout compact dans l'ouvert  $\Omega \setminus P$ , la fonction  $f$  est holomorphe sur  $\Omega \setminus P$  et la formule (26.370) est valable sur  $\Omega \setminus P$ . □

## 26.7 Dénombrement des solutions d'une équation diophantienne

Le théorème 26.102 peut être vu soit comme un dénombrement de solutions d'une certaine équation diophantienne, soit comme partition d'un entier en parts fixées. Avant de nous lancer dans sa démonstration, nous prouvons un certain nombre de lemmes qui vont traiter des aspects combinatoires de la preuve.

Soit  $n, N \in \mathbb{N}$ . Soit  $a \in \mathbb{N}^N$ . Nous considérons les ensembles suivants :

$$V_n(N) = \{x \in \mathbb{N}^N \text{ tel que } \sum_{i=1}^N x_i = n\}, \quad (26.371a)$$

$$W_n(a, N) = \{y \in \mathbb{N}^N \text{ tel que } y \cdot a = n\}, \quad (26.371b)$$

$$V_n(N)_a = \{x \in V_n(N) \text{ tel que } a_i \mid x_i \forall i = 1, \dots, N\}. \quad (26.371c)$$

L'ensemble  $V_n(N)$  avait déjà été rencontré en (15.67).

### Lemme 26.97.

*L'application*

$$\begin{aligned} \psi: W_n(a, N) &\rightarrow V_n(N)_a \\ (y_1, \dots, y_N) &\mapsto (y_1 a_1, \dots, y_N a_N) \end{aligned} \quad (26.372)$$

(1) *est bien définie, c'est-à-dire qu'elle prend effectivement ses valeurs dans  $V_n(N)_a$ ,*

(2) *est une bijection.*

*Démonstration.* En plusieurs parties.

- (1) **Bien définie** Soit  $y \in W_n(a, N)$ . Nous avons  $\psi(y)_i = a_i y_i$ . Le nombre  $a_i$  divise donc bien  $\psi(y)_i$  et  $\sum_{i=1}^N \psi(y)_i = n$ .
- (2) **Injective** Si  $\psi(y) = \psi(y')$ , alors pour tout  $i$  nous avons  $y_i a_i = y'_i a_i$ , et donc  $y_i = y'_i$ . La fonction  $\psi$  est donc bien injective.
- (3) **Surjective** Soit  $x \in V_n(N)_a$ . Vu que  $a_i \mid x_i$ , il existe  $y_i \in \mathbb{N}$  tel que  $x_i = a_i y_i$ . On vérifie que  $y \in W_n(a, N)$  et que  $\psi(y) = x$ .

□

LEMooOPXHooHzoHrm

### Lemme 26.98.

*Pour  $i = 1, \dots, N$ , nous posons*

$$\begin{aligned} b_i: \mathbb{N}^N &\rightarrow \{0, 1\} \\ x &\mapsto \begin{cases} 1 & \text{si } a_i \mid x_i \\ 0 & \text{sinon.} \end{cases} \end{aligned} \quad (26.373)$$

*Nous avons*

$$\sum_{x \in V_n(N)} \prod_{i=1}^N b_i(x) = \text{Card}(W_n(a, N)). \quad (26.374)$$

*Démonstration.* Nous décomposons la somme en  $V_n(N)_a$  et son complémentaire dans  $V_n(N)$  :

$$\sum_{x \in V_n(N)} \prod_{i=1}^N b_i(x) = \sum_{x \in V_n(N)_a} \prod_{i=1}^N b_i(x) + \sum_{x \in V_n(N) \setminus V_n(N)_a} \prod_{i=1}^N b_i(x) \quad (26.375a)$$

$$= \sum_{x \in V_n(N)_a} 1 + \sum_{x \in V_n(N) \setminus V_n(N)_a} 0 \quad (26.375b)$$

$$= \text{Card}(V_n(N)_a) \quad (26.375c)$$

$$= \text{Card}(W_n(N, a)) \quad (26.375d)$$

Justifications :

- Pour (26.375b). Si  $x \in V_n(N)_a$ , alors  $a_i \mid x_i$  pour tout  $i$ , et donc  $b_i(x) = 1$  pour tout  $i$ . Si au contraire  $x \in V_n(N) \setminus V_n(N)_a$ , il existe un  $i$  tel que  $a_i$  ne divise pas  $x_i$  et donc tel que  $b_i(x) = 0$ .
- Pour (26.375d). Les deux ensembles sont en bijection par le lemme 26.97.

□

LEMooRJOKooPJGVTr

**Lemme 26.99.**

Soit  $s \geq 1$ . La série entière

$$\sum_{n=0}^{\infty} z^{ns} \quad (26.376)$$

a un rayon de convergence égal à 1. Pour  $z \in B(0, 1)$  nous avons

$$\sum_{n=0}^{\infty} z^{ns} = \frac{1}{1 - z^s}. \quad (26.377)$$

EQooIRAZooKQooZnp

*Démonstration.* Pour un  $z \in \mathbb{C}$  fixé, nous avons  $z^{ns} = (z^s)^n$ <sup>58</sup>. Nous appelons donc la proposition 10.159 avec  $q = z^s$ .

Si  $|z| < 1$ , alors  $|z^s| < 1$  et la proposition 10.159 nous dit que la série converge. Si au contraire  $|z| > 1$ , alors  $|z^s| > 1$  et la série diverge.

Le corolaire 15.19 conclut que le rayon de convergence est bien 1.

La valeur (26.377) est également une partie de la proposition 10.159.

□

LEMooVMLEooCzPuKy

**Lemme 26.100.**

Si  $a, b \in \mathbb{C}$ , si  $N \in \mathbb{N}$  et si  $p \in \mathbb{N}$  avec  $p < N$ , nous posons

$$a_n = a \frac{(n + N - 1)!}{n!} \quad (26.378)$$

et

$$b_n = b \frac{(n + p - 1)!}{n!}. \quad (26.379)$$

Nous avons  $a_n \sim a_n + b_n$ .

*Démonstration.* Nous posons  $\alpha(n) = (a_n + b_n)/a_n$  et nous prouvons que  $\lim_{n \rightarrow \infty} \alpha(n) = 1$ . Pour

58. Vu qu'ici  $n$  et  $s$  sont entiers, c'est pas profond ça. Il ne faut pas invoquer la proposition générale 12.432.

ce faire,

$$\frac{b_n}{a_n} = \frac{b}{a} \frac{(n+p-1)!}{n!} \frac{n!}{(n+N-1)!} \quad (26.380a)$$

$$= \frac{b}{a} \frac{(n+p-1)!}{(n+N-1)!} \quad (26.380b)$$

$$= \frac{b}{a} \frac{(n+p-1)!}{(n+p-1)! \prod_{k=n+p}^{n+N-1} k} \quad (26.380c)$$

$$= \frac{b}{a} \prod_{k=n+p}^{n+N-1} \frac{1}{k} \quad (26.380d)$$

$$\leq \frac{b}{a} \frac{1}{n+p}, \quad (26.380e)$$

et nous avons

$$\lim_{n \rightarrow \infty} \frac{b_n}{a_n} = 0, \quad (26.381)$$

de telle sorte que  $\lim_{n \rightarrow \infty} \alpha(n) = 1$ . □

LEMooTGHHooZHZsgE

**Lemme 26.101.**

Si  $N \in \mathbb{N}$ , nous avons équivalence des suites

$$\frac{(n+N-1)!}{n!} \sim n^{N-1}. \quad (26.382)$$

*Démonstration.* Sachez que dans  $\prod_{i=a}^b$ , il y a  $b-a+1$  facteurs, et non  $b-a$  comme on pourrait naïvement le croire. Cela dit, nous avons le calcul

$$\frac{(n+N-1)!}{n!n^{N-1}} = \left( \prod_{k=n+1}^{n+N-1} k \right) \frac{1}{n^{N-1}} \quad (26.383a)$$

$$= \prod_{k=n+1}^{n+N-1} \frac{k}{n} \quad (26.383b)$$

$$= \prod_{k=1}^{N-1} \frac{n+k}{n}. \quad (26.383c)$$

Donc la limite

$$\lim_{n \rightarrow \infty} \frac{(n+N-1)!}{n!n^{N-1}} = 1. \quad (26.384)$$

□

THOooQDYWooCOiUMb

**Théorème 26.102** ([1, 136, 486, 707, 126]).

Soient  $N \in \mathbb{N}$  et  $a \in \mathbb{N}^N$  tel que  $\text{pgcd}(a_1, \dots, a_N) = 1$ . Nous posons

$$W_n(a, N) = \{y \in \mathbb{N}^N \text{ tel que } y \cdot a = n\}. \quad (26.385)$$

Nous avons alors<sup>59</sup>

$$\text{Card}(W_n(a, N)) \sim \frac{1}{\prod_{k=1}^N a_k} \frac{n^{N-1}}{(N-1)!}. \quad (26.386)$$

*Démonstration.* Pour chaque  $i = 1, \dots, N$ , nous considérons la série entière

$$s_i(z) = \sum_{k=0}^{\infty} z^{ka_i}. \quad (26.387)$$

---

59. Équivalence de suites, définition 9.33.

dont le rayon de convergence vaut 1 par le lemme 26.99. Nous nous apprêtons à faire le produit de Cauchy multiple de la proposition 15.31 ; nous posons donc

$$b_{ik} = \begin{cases} 1 & \text{si } a_i | k \\ 0 & \text{sinon,} \end{cases} \quad (26.388)$$

et nous écrivons toutes les séries  $s_i$  sous la forme

$$s_i(z) = \sum_{k=0}^{\infty} b_{ik} z^k. \quad (26.389)$$

La proposition 15.31 nous assure que si  $|z| < 1$ , le produit  $\prod_{i=1}^N s_i(z)$  peut être écrit sous la forme de la série entière

$$\prod_{i=1}^N s_i(z) = \sum_{s=0}^{\infty} \left( \sum_{x \in V_s(N)} \prod_{i=1}^N b_{ix_i} \right) z^s \quad (26.390a)$$

$$= \sum_{s=0}^{\infty} \left( \sum_{x \in V_s(N)} \prod_{i=1}^N b_i(x) \right) z^s \quad (26.390b)$$

$$= \sum_{s=0}^{\infty} \text{Card}(W_s(a, N)) z^s \quad (26.390c)$$

Justifications :

— Pour (26.390b). Notation  $b_i$  du lemme 26.98.

— Pour (26.390c). Utilisation du lemme 26.98

D'autre part, le lemme 26.99 nous permet d'écrire

$$f(z) = \prod_{i=1}^N s_i(z) = \prod_{i=1}^N \frac{1}{1 - z^{a_i}}. \quad (26.391)$$

Notre but sera d'écrire ce produit sous forme de série entière est d'identifier les coefficients avec ceux que l'on trouve dans (26.390c).

(1)  $m_\omega = N$  si et seulement si  $\omega^{a_i} = 1$  Nous montrons à présent que  $\omega$  est un pôle d'ordre  $N$  de  $f$  si et seulement si  $\omega^{a_1} = \dots = \omega^{a_N} = 1$ .

(1a) Un polynôme Nous considérons le polynôme

$$P(X) = \prod_{i=1}^N (1 - X^{a_i}) = (-1)^N \prod_{i=1}^N \prod_{\omega \in U_{a_i}} (X - \omega) \quad (26.392)$$

où, pour la seconde égalité, nous avons utilisé le lemme 19.20. En posant  $U = \bigcup_{i=1}^N U_{a_i}$  nous écrivons encore

$$P(X) = (-1)^N \prod_{\omega \in U} (X - \omega)^{m_\omega} \quad (26.393)$$

où  $m_\omega = \text{Card}\{i \text{ tel que } \omega \in U_{a_i}\}$ . Cela pour dire que, pour  $|z| < 1$ ,

$$f(z) = \frac{1}{P(z)} = \frac{(-1)^N}{\prod_{\omega \in U} (z - \omega)^{m_\omega}}. \quad (26.394)$$

(1b) Sens  $\Rightarrow$  Si  $\omega$  est un pôle d'ordre  $N$ , alors  $N = m_\omega = \{i \text{ tel que } \omega \in U_{a_i}\}$ . Donc  $\omega \in U_{a_i}$  pour tout  $i$ , c'est-à-dire que  $\omega^{a_i} = 1$  pour tout  $i$ .

(1c) Sens  $\Leftarrow$  Dans l'autre sens, si  $\omega^{a_i} = 1$  pour tout  $i$ , alors  $m_\omega = N$  et  $\omega$  est un pôle d'ordre  $N$ .

Nous pouvons continuer.

- (2)  $m_\omega = N$  si et seulement si  $\omega = 1$  Nous savons que  $\omega = 1$  est un pôle d'ordre  $N$  parce que  $1 \in U_{a_i}$  pour tout  $i$ . Dans l'autre sens, si  $\omega$  est d'ordre  $N$ , alors nous venons de voir que  $\omega^{a_i} = 1$  pour tout  $i$ .

Vu que les  $a_i$  sont premiers entre eux, le théorème de Bézout 3.59 nous donne des entiers  $u_i$  tels que  $\sum_{i=1}^N u_i a_i = 1$ . Nous avons alors

$$\omega = \omega^{u \cdot a} = \prod_{i=1}^N (\omega^{a_i})^{u_i} = 1. \quad (26.395)$$

Donc les pôles de  $f$  différents de 1 sont d'ordre strictement inférieur à  $N$ .

- (3) Décomposition en éléments simples Décomposons un peu l'expression de  $f(z)$  en repartant de (26.392) :

$$f(z) = \prod_{i=1}^N \frac{1}{1 - z^{a_i}} \quad (26.396a)$$

$$= \prod_{i=1}^N \frac{1}{-\prod_{\omega \in U_{a_i}} (z - \omega)} \quad (26.396b)$$

$$= (-1)^N \prod_{i=1}^N \prod_{\omega \in U_{a_i}} \frac{1}{z - \omega} \quad (26.396c)$$

$$= (-1)^N \prod_{i=1}^N \sum_{\omega \in U_{a_i}} \frac{\lambda_{\omega,i}}{z - \omega} \quad (26.396d)$$

Justifications :

— Pour (26.396b). Lemme 19.20 ; vous noterez le signe de différence.

— Pour (26.396d). Lemme 19.15 pour la décomposition en éléments simples.

- (4) Isoler le terme  $\omega = 1$  Nous notons  $U = \bigcup_{i=1}^N U_{a_i}$ . Chaque  $U_{a_i}$  contient  $\omega = 1$ . L'expression (26.396d) contient donc un terme en  $\frac{1}{(z-1)^N}$ . Tous les autres  $\omega$  de  $U$  ne sont présents que dans au maximum  $N-1$  des  $U_{a_i}$ . Nous avons donc

$$f(z) = \frac{A}{(z-1)^N} + \sum_{p=1}^{N-1} \sum_{\omega \in U} \frac{B_{p,\omega}}{(z-\omega)^p} \quad (26.397)$$

avec  $B_{p,\omega} \in \mathbb{C}$ .

- (5) Une belle lampée de factorielles Le lemme 15.53(3) permet d'écrire  $f$  avec des séries entières :

$$\begin{aligned} f(z) &= \frac{A}{(N-1)!} \sum_{s=0}^{\infty} \frac{(s+N-1)!}{s!} z^s \\ &+ \sum_{p=1}^{N-1} \sum_{\omega \in U} \frac{(-1)^p B_{p,\omega}}{(p-1)!} \sum_{s=0}^{\infty} \frac{(s+p-1)!}{s!} \frac{z^s}{\omega^{s+p+1}} \end{aligned} \quad (26.398)$$

qui est valable pour  $z \in B(0, 1)$ .

- (6) Ce qu'on en fait Pour rappel, l'équation (26.390) nous dit que

$$f(z) = \sum_{s=0}^{\infty} \text{Card}(W_s(a, N)) z^s. \quad (26.399)$$

Nous allons donc identifier le coefficient de  $z^n$  dans (26.398) avec  $\text{Card}(W_n(a, N))$  :

$$\begin{aligned} \text{Card}(W_n(a, N)) &= \frac{A}{(N-1)!} \frac{(n+N-1)!}{n!} \\ &+ \sum_{p=1}^{N-1} \sum_{\omega \in U} \frac{(-1)^p B_{p,\omega}}{(p-1)!} \frac{(n+p-1)!}{n!} \frac{1}{\omega^{n+p+1}}. \end{aligned} \quad (26.400)$$

Voici une belle suite (par rapport à  $n$ ) dont nous devons étudier le comportement asymptotique.

- (7) **Des équivalences** Le lemme 26.100 nous permet de supprimer tous les termes autres que celui qui contient  $A$  :

$$\text{Card}(W_n(a, N)) \sim \frac{A}{(N-1)!} \frac{(n+N-1)!}{n!}. \quad (26.401)$$

Notez que, à gauche, nous avons une suite dans  $\mathbb{N}$  et à droite, une suite dans  $\mathbb{C}$  (il n'y a pas de raisons à priori que  $A$  soit entier ou réel). Cela n'a pas d'importance ; ça n'empêche pas les suites d'être équivalentes.

Le lemme 26.101 donne maintenant

$$\text{Card}(W_n(a, N)) \sim \frac{A}{(N-1)!} n^{N-1}. \quad (26.402) \quad \text{EQooTPXCooGHSzoP}$$

Cela est déjà très bien parce que ça donne la vitesse de croissance en fonction de  $N$  et  $n$ . Mais puisque nous sommes perfectionnistes, nous allons encore déterminer la valeur de  $A$ .

- (8) **La valeur de  $A$**  Pour déterminer la valeur de  $A$ , l'astuce est de considérer la fonction  $z \mapsto f(z)(1-z)^N$  :

$$f(z)(1-z)^N = (1-z)^N \prod_{i=1}^N \frac{1}{1-z^{a_i}} \quad (26.403a)$$

$$= \prod_{i=1}^N \frac{1-z}{1-z^{a_i}} \quad (26.403b)$$

$$= \prod_{i=1}^N \frac{1}{1+\dots+z^{a_i-1}}. \quad (26.403c) \quad \text{SUBEQooUZTCooYgHaES}$$

Justifications :

— Pour (26.403c). C'est le lemme 3.124(2).

La dernière expression montre qu'il n'y a pas de mal à prendre la limite  $z \rightarrow 1$  ; elle vaut

$$\lim_{z \rightarrow 1} f(z)(1-z)^N = \prod_{i=1}^N \frac{1}{a_i}. \quad (26.404)$$

Mais en partant d'autre part de (26.397), nous avons

$$f(z)(1-z)^N = A + \sum_{p=1}^{N-1} \sum_{\omega \in U} \frac{B_{p,\omega}(1-z)^N}{(z-\omega)^p}. \quad (26.405)$$

Vu que  $N > p$ , la limite  $z \rightarrow 0$  existe et vaut zéro dans tous les éléments de la somme, y compris les éléments avec  $\omega = 1$ . Donc

$$\lim_{z \rightarrow 1} f(z)(1-z)^N = A. \quad (26.406)$$

Nous savons donc que

$$A = \prod_{i=1}^N \frac{1}{a_i}. \quad (26.407) \quad \text{EQooJMALooUrXJZc}$$

En remettant la valeur (26.407) dans l'équivalence (26.402), nous trouvons le résultat demandé.  $\square$

### Exemple 26.103.

Pour  $p = 1$ , l'équation est  $\alpha x = n$ , qui possède au maximum une solution, quel que soit  $n$ . Et de plus pour avoir une solution il faut et suffit que  $\alpha$  divise  $n$ , c'est-à-dire que  $n$  soit un multiple de  $\alpha$ . Il n'y a que un nombre sur  $\alpha$  à être multiple de  $\alpha$ . D'où le comportement en  $\frac{1}{\alpha}$ .

Pour  $p = 2$ , c'est l'équation (3.53) déjà étudiée. Il y a une famille à un paramètre de solutions dont seulement un certain nombre sont positives. À priori, le nombre de solutions positives croît linéairement en  $n$ .  $\triangle$



**26.104.**

Si vous aimez les séries génératrices. Si vous aimez l'idée de mettre toute l'information d'un problème dans les coefficients d'une série puis de trouver des réponses en les manipulant, vous pouvez regarder *introduction à la théorie analytique des nombres*[708].

Cette vidéo explique comment payer  $n$  euros avec des pièces et des billets de valeur données. On pourrait croire que cela est exactement le résultat du théorème 26.102. Il n'en est rien parce que l'hypothèse de pgcd du théorème n'est pas du tout réalisée par les pièces et billets actuellement en circulation.

Du coup, je ne sais pas si ce théorème est intéressant au sens de la définition 0.2.

**26.8 Fonctions d'Euler**

ThoZJYooWKfbVz

**Théorème 26.105** (Prolongement méromorphe de la fonction  $\Gamma$  d'Euler[126]).

Nous considérons la formule

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt. \quad (26.408)$$

Alors

(1) Cette formule définit une fonction holomorphe sur

$$\mathcal{P} = \{z \in \mathbb{C} \text{ tel que } \Re(z) > 0\}. \quad (26.409)$$

(2) La fonction  $\Gamma: \mathcal{P} \rightarrow \mathbb{C}$  admet un unique prolongement méromorphe sur  $\mathbb{C}$ , lequel a des pôles sur les entiers négatifs.

*Démonstration.* (1) **Holomorphie sous l'intégrale**

Pour étudier l'holomorphie de la fonction  $\Gamma$  sur  $\mathcal{P}$  nous utilisons le théorème 26.73.

Nous considérons la fonction

$$\begin{aligned} g: \mathcal{P} \times \mathbb{R}^+ &\rightarrow \mathbb{C} \\ (z, t) &\mapsto e^{-t} t^{z-1} \end{aligned} \quad (26.410)$$

et nous commençons par montrer que c'est holomorphe en  $z$  pour chaque  $t > 0$  fixé. Nous le vérifions par le critère de  $\partial_{\bar{z}} f = 0$ <sup>60</sup> et en nous souvenant que  $t^i = e^{\ln(t)i} = e^{i \ln(t)}$ . Nous obtenons rapidement que

$$\frac{\partial g}{\partial \bar{z}} = 0. \quad (26.411)$$

Le fait que la fonction  $t \mapsto g(z, t)$  soit mesurable pour tout  $z$  est d'accord.

Et enfin soit  $K$  compact dans  $\mathcal{P}$ . Il faut trouver une fonction  $g_K(t)$  intégrable sur  $[0, \infty[$  telle que pour tout  $z \in K$  et  $t \in [0, \infty[$  nous ayons  $|g(z, t)| \leq g_K(t)$ . Pour cela nous majorons séparément les parties  $t \in ]0, 1[$  et  $t \geq 1$ .

Soit donc  $K$  compact dans  $\mathcal{P}$ ; nous posons  $M = \max_{z \in K} \Re(z)$  et  $\epsilon = \min_{z \in K} \Re(z)$ .

Si  $t \in ]0, 1[$  alors nous avons

$$e^{-t} t^{z-1} = e^{-t} e^{(z-1) \ln(t)}, \quad (26.412)$$

de telle façon à que que

$$|e^{-t} t^{z-1}| \leq |e^{(x-1+iy) \ln(t)}| \quad (26.413a)$$

$$= |e^{(\Re(z)-1) \ln(t)}| \quad (26.413b)$$

$$= |t^{\Re(z)-1}| \quad (26.413c)$$

$$\leq |t^{\epsilon-1}| \quad (26.413d)$$

$$= \frac{1}{t^{1-\epsilon}}. \quad (26.413e)$$

---

60. Théorème 26.2.

Cette dernière fonction est intégrable sur  $]0, 1[$ .

Nous considérons maintenant  $t \geq 1$ . Dans ce cas nous avons

$$|e^{-t}t^{z-1}| = e^{-t}t^{\Re(z)-1} \leq e^{-t}t^{M-1}. \quad (26.414)$$

Cette dernière fonction est un produit d'une exponentielle décroissante avec un polynôme. C'est donc intégrable entre 1 et l'infini.

La fonction  $g_K$  que nous considérons est donc

$$g_K(t) = \begin{cases} \frac{1}{t^{1-\epsilon}} & \text{si } t < 1 \\ \text{borné} & \text{si } 1 \leq t \leq b \\ e^{-t}t^{M-1} & \text{si } t > b. \end{cases} \quad (26.415)$$

Cela est une fonction intégrable sur  $]0, \infty[$  et qui majore  $g$  uniformément en  $z$  sur le compact  $K$  de  $\mathcal{P}$ . Le théorème 26.73 nous permet donc de conclure que

$$\Gamma(z) = \int_0^\infty g(z, t) dt \quad (26.416)$$

est holomorphe en  $z$  sur  $\mathcal{P}$  et que

$$\Gamma'(z) = \int_0^\infty \frac{\partial g}{\partial z}(z, t) dt. \quad (26.417)$$

- (2) **En deux morceaux** Nous passons maintenant à la seconde partie du théorème. Pour  $z \in \mathcal{P}$  nous coupons l'intégrale en deux :

$$\Gamma(z) = \int_0^1 e^{-t}t^{z-1} dt + \int_1^\infty e^{-t}t^{z-1} dt \quad (26.418)$$

- (3) **Première partie** Nous commençons par parler de la première partie :  $\int_0^1 e^{-t}t^{z-1} dt$  dans laquelle nous voulons utiliser le développement en série de l'exponentielle  $e^{-t}$ . Nous devons donc traiter

$$\int_0^1 \sum_{n=0}^\infty \frac{(-1)^n}{n!} t^{n+z-1} dt. \quad (26.419)$$

Nous allons permuter la somme avec l'intégrale à l'aide du théorème de Fubini 14.305 en posant la fonction

$$g(n, t) = \frac{(-1)^n}{n!} t^{n+z-1} \quad (26.420)$$

et en considérant le produit entre la mesure de Lebesgue sur  $\mathbb{C}$  et la mesure de comptage sur  $\mathbb{N}$ , c'est-à-dire que nous étudions

$$\int_0^1 \int_{\mathbb{N}} g(n, t) dndt. \quad (26.421)$$

Pour permuter il suffit de prouver que  $|g|$  est intégrable pour la mesure produit, c'est-à-dire que

$$\int_0^1 \int_{\mathbb{N}} \left| \frac{(-1)^n}{n!} t^{n+z-1} \right| < \infty. \quad (26.422)$$

Nous avons  $|t^z| = t^{\Re(z)}$ , donc

$$\sum_{n=0}^\infty \left| \frac{t^{n+z-1}}{n!} \right| = t^{\Re(z)-1} \sum_{n=0}^\infty \frac{t^n}{n!} = t^{\Re(z)-1} e^t. \quad (26.423)$$

Étant donné que nous avons fixé  $z \in \mathcal{P}$ , nous avons  $\Re(z) - 1 > -1$  et donc  $t^{\Re(z)-1}$  est intégrable entre 0 et 1 (proposition 14.301). La partie  $e^t$  se majore sur  $[0, 1]$  par une constante quelconque. Nous avons donc payé le droit d'inverser la somme et l'intégrale :

$$\int_0^1 e^{-t} t^{z-1} dt = \sum_{n=0}^{\infty} \int_0^1 \frac{(-1)^n}{n!} t^{n+z-1} dt = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} [t^{n+z}]_0^1 = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+z)}. \quad (26.424)$$

Nous avons donc l'intéressante formule suivante, valable pour tout  $z \in \mathcal{P}$  :

$$\Gamma(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+z)} + \int_1^{\infty} e^{-t} t^{z-1} dt. \quad (26.425)$$

(4) **Prolongation de la première partie** Nous voudrions montrer maintenant que la fonction

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n-z)} \quad (26.426)$$

est méromorphe sur  $\mathbb{C}$  avec des pôles en les entiers négatifs. Pour cela nous considérons la suite de fonctions

$$f_n(z) = \frac{(-1)^n}{n!(z+n)} \quad (26.427)$$

et nous allons utiliser la proposition 26.96. Si  $n \geq 0$ , la fonction  $f_n$  est méromorphe sur  $\mathbb{C}$  avec un pôle simple en  $z = -n$ . Soit  $K$  compact de  $\mathbb{C}$  et  $N_K$  tel que  $K \subset \overline{B(0, N_K)}$ . Pour  $n \geq N_K + 1$ , la fonction  $f_n$  n'a pas de pôle dans  $K$  et de plus pour tout  $z \in K$  nous avons

$$|z+n| = |z-(-n)| \geq |n-|z|| \geq n-|z| \geq n-N_K, \quad (26.428)$$

et par conséquent

$$|f_n(z)| \leq \frac{1}{n!(n-N)}, \quad (26.429)$$

ou pour le dire de façon plus snob :

$$\|f_n\|_{\infty, K} \leq \frac{1}{n!(n-N)}, \quad (26.430)$$

dont la série converge. Cela signifie que la série  $\sum_{n \geq N} f_n$  converge normalement<sup>61</sup> sur  $K$ , donc la fonction

$$f(z) = \sum_{n=0}^{\infty} f_n(z) \quad (26.431)$$

est une fonction méromorphe dont les pôles sont ceux des  $f_n$ , c'est-à-dire les entiers négatifs (proposition 26.96).

(5) **La seconde partie**

Nous allons à présent prouver que la fonction

$$h(z) = \int_1^{\infty} e^{-t} t^{z-1} dt \quad (26.432)$$

est holomorphe sur  $\mathbb{C}$ . Pour cela nous considérons la fonction de deux variables  $f(z, t) = e^{-t} t^{z-1}$  et nous utilisons le théorème d'holomorphicité sous l'intégrale 26.73. D'abord pour  $z_0$  fixé dans  $\mathbb{C}$  nous avons

$$\int_1^{\infty} |e^{-t} t^{z_0-1}| \leq \int_1^{\infty} e^{-t} t^{\Re(z_0)-1} dt, \quad (26.433)$$

61. Définition 10.110.

donc l'intégrale converge parce que c'est polynôme contre exponentielle. Par ailleurs pour chaque  $t_0$  fixé sur  $[0, \infty[$ , la fonction  $z \mapsto e^{-t_0} t_0^{z-1}$  est holomorphe sur  $\mathbb{C}$  comme en témoigne le calcul suivant :

$$\frac{1}{2} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) t_0^{x+iy-1} = 0. \quad (26.434)$$

Et enfin si  $K$  est compact dans  $\mathbb{C}$  nous avons

$$|f(z, t)| = |e^{-t} t^{z-1}| = e^{-t} |t^{\Re(z)-1}| \leq e^{-t} t^{M-1} \quad (26.435)$$

où  $M = \max_{z \in K} \Re(z)$ . Nous en déduisons que la fonction

$$z \mapsto \int_1^\infty e^{-t} t^{z-1} dt \quad (26.436)$$

est une fonction holomorphe sur  $\mathbb{C}$ .

#### (6) **Conclusion**

Au final nous avons prouvé que la fonction  $\Gamma$  d'Euler admet le prolongement méromorphe sur  $\mathbb{C}$  donné par

$$\Gamma(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(z+n)} + \int_1^\infty e^{-t} t^{z-1} dt. \quad (26.437)$$

□

### 26.8.1 Euler et factorielle

#### **Proposition 26.106.**

Nous avons la formule  $\Gamma(n) = (n-1)!$  pour tout  $n \in \mathbb{N}$ .

*Démonstration.* Nous partons de la formule

$$\Gamma(n) = \int_0^\infty e^{-t} t^{n-1} dt \quad (26.438)$$

que nous intégrons par partie en posant

$$\begin{aligned} u &= t^{n-1} & u' &= (n-1)t^{n-2} \\ v &= e^{-t} & v' &= -e^{-t}. \end{aligned} \quad (26.439)$$

Les termes au bord s'annulent (ici il y a un passage à la limite qui n'est pas écrit) et nous trouvons

$$\Gamma(n) = \int_0^\infty (n-1)e^{-t} t^{n-2} dt = (n-1)\Gamma(n-1). \quad (26.440)$$

Pour conclure il suffit de remarquer que

$$\Gamma(1) = \int_0^\infty e^{-t} dt = -[e^{-t}]_0^\infty = 1. \quad (26.441)$$

□

## 26.9 Exponentielle et logarithme complexe

### 26.9.1 Propriétés de l'exponentielle

#### **Proposition 26.107.**

Soit  $z \in \mathbb{C}$  fixé. La fonction

$$\begin{aligned} E: \mathbb{R} &\rightarrow \mathbb{C} \\ t &\mapsto e^{tz} \end{aligned} \quad (26.442)$$

est  $C^\infty$ , sa dérivée est

$$E'(t) = ze^{tz}. \quad (26.443)$$

La fonction  $E$  est développable en série entière (voir définition 15.149) sur  $\mathbb{R}$  en  $t = 0$  et

$$e^{tz} = \sum_{n=0}^{\infty} \frac{z^n}{n!} t^n. \quad (26.444)$$

*Démonstration.* Nous fixons  $z \in \mathbb{C}$ . Par définition 18.6, la série suivante est  $e^{tz}$  :

$$f(t) = \sum_{n=0}^{\infty} \frac{z^n}{n!} t^n. \quad (26.445)$$

Cette série a un rayon de convergence infini et la fonction  $f$  est donc  $C^\infty$  sur  $\mathbb{R}$ . Nous pouvons la dériver terme à terme :

$$f'(t) = \sum_{n=1}^{\infty} \frac{z^n}{n!} n t^{n-1} = z \sum_{n=1}^{\infty} \frac{z^{n-1}}{(n-1)!} t^{n-1} = ze^{tz}. \quad (26.446)$$

□

THOooNGOIooEECFav

### **Théorème 26.108.**

La fonction exponentielle vérifie les propriétés suivantes.

- (1)  $\exp$  est holomorphe<sup>62</sup>.
- (2)  $(e^z)' = e^z$ .
- (3) L'exponentielle est développable en série entière,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad (26.447)$$

et la série converge normalement sur tout compact de  $\mathbb{C}$ .

*Démonstration.* En tant que application  $E: \mathbb{R}^2 \rightarrow \mathbb{C}$ , la fonction

$$E(x, y) = e^x (\cos y + i \sin y) \quad (26.448)$$

est  $C^\infty$ . De plus nous avons

$$\frac{\partial E}{\partial x}(x, y) = e^{x+iy} = E(x, y) \quad (26.449a)$$

$$\frac{\partial E}{\partial y}(x, y) = iE(x, y), \quad (26.449b)$$

et par conséquent la fonction  $E$  vérifie les équations de Cauchy-Riemann.

Si  $r$  est fixé, par le critère d'Abel appliqué à la suite  $r^n/n!$  nous savons que la série  $\sum z^n/n!$  converge normalement sur le compact  $B(0, r)$ . □

### **26.9.2 Intégrale de Fresnel**

Nous allons calculer l'intégrale de Fresnel

$$\int_0^\infty e^{-ix^2} dx = \frac{\sqrt{\pi}}{2} e^{-i\pi/4} \quad (26.450)$$

en suivant la démarche présentée par Wikipédia[709]. Nous commençons par prouver que l'intégrale est convergente en nous contentant de justifier la convergence de

$$\int_0^\infty \sin(x^2) dx. \quad (26.451)$$

---

62. Définition 12.331.

Pour chaque  $a > 0$  fixé, l'intégrale  $\int_0^a \sin(x^2) dx$  ne pose pas de problèmes. Le lemme 20.197 nous permet de passer à la limite ; nous devons donc seulement calculer

$$\lim_{b \rightarrow \infty} \int_a^b \sin(x^2) dx \quad (26.452)$$

où  $a$  est une constante strictement positive. Nous effectuons une intégration par partie en posant

$$u = \frac{1}{x} \quad u' = -\frac{1}{x^2} \quad (26.453a)$$

$$v' = x \sin(x^2) \quad v = \frac{1 - \cos(x^2)}{2}. \quad (26.453b)$$

Notons que la primitive  $v$  a été choisie pour avoir  $v(0) = 0$ . Nous avons

$$\int_a^b \sin(x^2) dx = \left[ \frac{1 - \cos(x^2)}{2x} \right]_a^b - \int_a^b \frac{\cos(x^2) - 1}{2x^2} dx \quad (26.454)$$

Pour le premier terme nous avons

$$\lim_{b \rightarrow \infty} \left[ \frac{1 - \cos(x^2)}{2x} \right]_a^b = \lim_{b \rightarrow \infty} \frac{1 - \cos(b^2)}{2b} - \frac{1 - \cos(a^2)}{2a} = -\frac{1 - \cos(a^2)}{2a}. \quad (26.455)$$

C'est borné. Pour le second terme de (26.454), la fonction

$$\frac{\cos(x^2) - 1}{2x^2} \quad (26.456)$$

est majorée par la fonction  $1/x^2$  qui est intégrable entre  $a$  et  $\infty$ .

Nous allons calculer l'intégrale demandée en passant par la fonction

$$f(x) = e^{-x^2} \quad (26.457)$$

définie sur le plan complexe. Nous l'intégrons sur le chemin  $\gamma = \gamma_1 + \gamma_2 - \gamma_3$  indiqué à la figure 26.1.

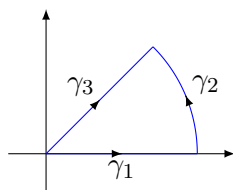


FIGURE 26.1: Chemin d'intégration pour l'intégrale de Fresnel LabelFigCheminFresnel

Ces chemins sont donnés par

$$\begin{aligned} \gamma_1: [0, R] &\rightarrow \mathbb{C} \\ t &\mapsto t, \end{aligned} \quad (26.458)$$

$$\begin{aligned} \gamma_2: [0, \frac{\pi}{4}] &\rightarrow \mathbb{C} \\ t &\mapsto Re^{it}, \end{aligned} \quad (26.459)$$

$$\begin{aligned} \gamma_3: [0, R] &\rightarrow \mathbb{C} \\ t &\mapsto te^{i\pi/4}. \end{aligned} \quad (26.460)$$

Tout d'abord la fonction  $f$  est bien holomorphe par le critère du théorème 26.2. Le calcul de  $\frac{\partial f}{\partial \bar{z}}$  se fait simplement en posant  $f(x, y) = e^{-(x+iy)^2}$ . Le calcul est usuel :

---

```
| Sage Version 4.8, Release Date: 2012-01-20 |
| Type notebook() for the GUI, and license() for information. |
```

---

```
sage: f(x,y)=exp(-(x+I*y)**2)
sage: A=f.diff(x)+I*f.diff(y)
sage: A.simplify_full()
(x, y) |--> 0
```

Nous avons donc

$$0 = \int_{\gamma} f = \underbrace{\int_0^R e^{-t^2} dt}_{I_1(R)} + \underbrace{\int_0^{\pi/4} e^{-R^2 e^{2it}} R i e^{it} dt}_{I_2(R)} + \underbrace{\int_0^R e^{-t^2 e^{i\pi/2}} e^{i\pi/4} dt}_{I_3(R)}. \quad (26.461)$$

L'intégrale est nulle pour tout  $R$  en vertu de la proposition 26.11. L'intégrale  $I_1$  est une gaussienne et nous avons

$$\lim_{R \rightarrow \infty} I_1(R) = \frac{\sqrt{\pi}}{2} \quad (26.462)$$

par l'exemple 14.308. Nous montrons maintenant que  $\lim_{R \rightarrow \infty} |I_2(R)| = 0$ <sup>63</sup>. D'abord nous majorons en prenant la norme puis nous effectuons le changement de variables  $u = 2t$  :

$$|I_2(R)| \leq \int_0^{\pi/4} R e^{-R^2 \cos(2t)} dt \quad (26.463a)$$

$$= \frac{R}{2} \int_0^{\pi/2} e^{-R^2 \cos(u)} du. \quad (26.463b)$$

Nous savons que le graphe du cosinus est concave : il reste au dessus de la droite que joint  $(0, 1)$  à  $(\frac{\pi}{2}, 0)$ . Du coup  $\cos(u) \geq 1 - \frac{2}{\pi}u$  et par conséquent

$$e^{-R^2 \cos(u)} \leq e^{-R^2(1 - \frac{2}{\pi}u)} = e^{R^2(\frac{2}{\pi}u - 1)}. \quad (26.464)$$

Nous effectuons l'intégrale

$$|I_2(R)| \leq \frac{R}{2} \int_0^{\pi/2} e^{-R^2} e^{\frac{2R^2}{\pi}u} du \quad (26.465a)$$

$$= \frac{R}{2} e^{-R^2} \left[ \frac{\pi}{2R^2} e^{2R^2 u/\pi} \right]_0^{\pi/2} \quad (26.465b)$$

$$= \frac{\pi}{4R} - \frac{\pi e^{-R^2}}{4R}, \quad (26.465c)$$

et nous avons bien  $\lim_{R \rightarrow \infty} |I_2(R)| = 0$ . Nous passons à la troisième intégrale. En tenant compte que  $e^{i\pi/2} = i$ , nous avons

$$I_3(R) = - \int_0^R e^{-\gamma_3(t)^2} e^{i\pi/4} dt \quad (26.466a)$$

$$= - \frac{1+i}{\sqrt{2}} \int_0^R e^{-t^2} e^{2i\pi/4} dt \quad (26.466b)$$

$$= - \frac{1+i}{\sqrt{2}} \int_0^R e^{-it^2} dt. \quad (26.466c)$$

En passant à la limite  $R \rightarrow \infty$ , de l'équation (26.461) il ne reste que

$$0 = \frac{\sqrt{2}}{2} - \frac{1+i}{\sqrt{2}} \int_0^{\infty} e^{-it^2} dt, \quad (26.467)$$

ce qui signifie que

$$\int_0^{\infty} e^{-it^2} dt = \frac{\sqrt{2\pi}}{2(1+i)} = \frac{\sqrt{\pi}}{2} e^{-i\pi/4}. \quad (26.468)$$

---

<sup>63</sup>. Il y a moyen de démontrer cela via le lemme de Jordan[710]. Nous donnons ici une démonstration moins technologique.

## 26.10 Théorème de Weierstrass

ThoArYtQO

**Théorème 26.109** (Théorème de Weierstrass[711]).

Soit  $(f_n)$  une suite de fonctions holomorphes sur un ouvert  $\Omega$  de  $\mathbb{C}$  que nous supposons converger uniformément sur tout compact vers  $f$ . Alors  $f$  est holomorphe sur  $\Omega$  et pour tout  $k$  nous avons

$$f_n^{(k)} \rightarrow f^{(k)} \quad (26.469)$$

uniformément sur tout compact.

Dit en peu de mots, la limite uniforme d'une suite de fonctions holomorphes est holomorphe, et on peut permuter la limite avec la dérivation.

*Démonstration.* Chacune des fonctions  $f_n$  étant holomorphes, si  $a \in \Omega$  et  $r$  est tel que  $B(a, r) \subset \Omega$ , nous avons par la formule de Cauchy 26.60 :

$$f_n(z) = \frac{1}{2\pi i} \int_{\partial B(a, r)} \frac{f_n(\xi)}{\xi - z} d\xi \quad (26.470)$$

pour tout  $z$  dans un boule  $B(a, \rho)$  incluse dans  $B(a, r)$ . Étant donné que le cercle  $\partial B$  est compact, elle y est majorée par une constante  $M$ . Montrons que de plus nous pouvons choisir  $M$  de telle façon à avoir  $|f_n(\xi)| \leq M$  pour tout  $n$  et tout  $\xi$  en même temps. D'abord nous utilisons la continuité de la limite  $f$  sur le compact  $\partial B$  pour poser  $A = \max_{z \in \partial B} |f(z)|$ . Ensuite nous considérons un  $\epsilon > 0$  et  $N$  tel que  $\|f_n - f\|_{\partial B} \leq \epsilon$  pour tout  $n \geq N$ . Nous savons maintenant que

$$\{|f_n(\xi)| \text{ tel que } n \geq N, \xi \in \partial B\} \quad (26.471)$$

est majoré par  $A + \epsilon$ . Nous posons enfin

$$B = \max_{n \leq N} \max_{\xi \in \partial B} |f_n(z)|, \quad (26.472)$$

et alors le nombre  $M = \max\{A + \epsilon, B\}$  majore  $|f_n(\xi)|$  pour tout  $n$  et tout  $\xi \in \partial B$ .

De plus pour tout  $\xi \in \partial B$  et pour tout  $z$  dans la petite boule, nous avons  $|\xi - z| > r - \rho$ , donc la fonction dans l'intégrale est majorée par une constante ne dépendant ni de  $n$  ni de  $\xi$ . Nous pouvons donc permuter l'intégrale et la limite sur  $n$  :

$$f(z) = \frac{1}{2i\pi} \int_{\partial B} \frac{f(\xi)}{\xi - z}. \quad (26.473)$$

Cela implique que la fonction  $f$  est holomorphe par le corolaire 26.63.

Nous voudrions maintenant parler des dérivées des  $f_n$  et de  $f$ . Pour cela nous voulons permuter l'intégrale et les dérivées, ce qui est fait au corolaire 26.75 :

$$f_n^{(k)} = \frac{1}{2\pi i} \int_{\partial B(z_0, r)} \frac{f(\omega)}{(\omega - z)^{k+1}} d\omega. \quad (26.474)$$

Nous voulons la convergence sur tout compact contenu dans l'ouvert  $\Omega$ . Pour ce faire, nous allons considérer un compact  $K \subset \Omega$  et prouver la convergence uniforme dans toute boule de la forme  $B(z_0, r)$  avec  $z_0 \in K$  et  $B(z_0, r) \subset \Omega$ . Pour chaque tel couple  $(z_0, r)$ , nous aurons un  $N_{(z_0, r)} \in \mathbb{N}$  tel que si  $n \geq N_{(z_0, r)}$ ,

$$\|f_n^{(k)} - f^{(k)}\|_{B(z_0, r)} \leq \epsilon. \quad (26.475)$$

Vu que ces boules  $B(z_0, r)$  forment un recouvrement de  $K$  par des ouverts, nous pouvons en retirer un sous-recouvrement fini et prendre, comme  $N$ , le maximum des  $N_{(z_0, r)}$  correspondants. Pour ce  $N$  nous aurons

$$\|f_n^{(k)} - f^{(k)}\|_K \leq \epsilon. \quad (26.476)$$

Au travail !



Pour  $z \in B(z_0, r)$  nous considérons  $r' > r$  tel que  $B(z_0, r') \subset \Omega$  et nous avons

$$|f_n^{(k)}(z) - f^{(k)}(z)| = \left| \frac{1}{2\pi i} \int_{\partial B(z_0, r')} \frac{f_n(\xi) - f(\xi)}{(\xi - z)^{k+1}} d\xi \right| \quad (26.477a)$$

$$\leq \frac{1}{2\pi} \int_{\partial B(z_0, r')} \frac{|f_n(\xi) - f(\xi)|}{|r - r'|^{k+1}} d\xi. \quad (26.477b)$$

Nous avons pris ce  $r'$  de telle manière que  $|\xi - z|$  soit borné par le bas par  $|r - r'|$ ; sinon la majoration que nous venons de faire ne marche pas. Étant donné que  $f_n \rightarrow f$  uniformément, nous pouvons considérer  $n$  assez grand pour que le numérateur soit plus petit que  $\epsilon$  indépendamment de  $\xi$  et de  $z$ . Donc pour un  $n$  assez grand,

$$|f_n^{(k)}(z) - f^{(k)}(z)| \leq \frac{\epsilon}{2\pi} \frac{2\pi r'}{|r - r'|^{k+1}} \quad (26.478)$$

pour tout  $z \in B(z_0, r)$ . Donc nous avons convergence uniforme  $f_n^{(k)} \rightarrow f^{(k)}$  sur cette boule. Par l'argument de compacité donné plus haut, nous avons la convergence uniforme sur tout compact.  $\square$

## 26.11 Analyse complexe en plusieurs variables

**Définition 26.110** ([712]).

Nous définissons le **polydisque** centré en  $a \in \mathbb{C}^n$  et de « rayons »  $r \in \mathbb{C}^n$  par

$$D(a, r) = \{z \in \mathbb{C}^n \text{ tel que } |z_j - a_j| < r_j \forall j = 1, \dots, n\}. \quad (26.479)$$

**Définition 26.111.**

Pour une fonction  $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$ , nous notons

$$\frac{\partial f}{\partial z_j} = \frac{1}{2} \left( \frac{\partial}{\partial x_j} - i \frac{\partial}{\partial y_j} \right) f. \quad (26.480)$$

**Définition 26.112** ([712]).

Soit un ouvert<sup>64</sup>  $D$  dans  $\mathbb{C}^n$ . Nous disons que  $f: D \rightarrow \mathbb{C}$  est **holomorphe** si

- (1)  $f$  est continue sur  $D$ ,
- (2) pour tout  $z \in D$  et pour tout  $1 \leq j \leq n$ , le nombre  $\frac{\partial f}{\partial z_j}(z)$  existe et est fini.

**Définition 26.113** ([713]).

Soient des ouverts  $B_1$  et  $B_2$  dans  $\mathbb{C}^n$ . Nous disons que  $f: B_1 \rightarrow B_2$  est **biholomorphe** si

- (1)  $f$  est holomorphe,
- (2)  $f$  est une bijection entre  $B_1$  et  $B_2$ ,
- (3)  $f^{-1}$  est holomorphe.

THOooNBGZooHuGtxW

**Théorème 26.114** (Inversion locale, version holomorphe).

Soient des ouverts  $B_1$  et  $B_2$  dans  $\mathbb{C}^n$ . Nous considérons une fonction holomorphe  $f: B_1 \rightarrow B_2$ . Soit  $z_0 \in B_1$  et  $w_0 = f(z_0)$ . Il y a équivalence entre

- (1)  $\det(df_{z_0}) \neq 0$
- (2) Il existe un voisinage  $U$  de  $z_0$  dans  $B_1$  et  $V$  de  $w_0$  dans  $B_2$  tels que la restriction  $f: U \rightarrow V$  est biholomorphe.

64. La topologie sur  $\mathbb{C}^n$  est celle de la norme définie en 9.117.

# Chapitre 27

## Analyse fonctionnelle

### 27.1 Théorème d'isomorphisme de Banach

ThofQShsw

**Théorème 27.1** (théorème d'isomorphisme de Banach[714]).

Une application linéaire continue et bijective entre deux espaces de Banach est un homéomorphisme<sup>1</sup>.

*Démonstration.* Soit une application linéaire bijective et continue  $f: E \rightarrow F$  entre deux espaces de Banach. En particulier elle est surjective, et le théorème de l'application ouverte 10.184 s'applique :  $f$  est une application ouverte.

Vu que  $f$  est bijective et ouverte, la proposition 10.180 implique que  $f^{-1}$  est continue. L'application  $f$  est donc continue d'inverse continue. Elle est donc un homéomorphisme.  $\square$

#### ⚠ Avertissement/question au lecteur !! 27.2

Il est donné dans [715] un exemple d'application linéaire et bijective dont l'inverse n'est pas continue. Si un jour vous le mettez au propre, faites-m'en profiter en m'envoyant une photo de votre feuille.

### 27.2 Théorème d'Ascoli

PropDGsPtpU

**Proposition-Définition 27.3** (Application compacte[466]).

Soient  $E$  et  $F$  deux espaces vectoriels normés sur  $\mathbb{R}$  ou  $\mathbb{C}$  et une application  $f \in \mathcal{L}(E, F)$ . Les propriétés suivantes sont équivalentes.

ITEMooAFNXooUGbZsh

(1) L'image d'un borné de  $E$  par  $f$  est relativement compacte<sup>2</sup> dans  $F$ .

ItemJlkpUblIi

(2) L'image par  $f$  de la boule unité fermée est relativement compacte dans  $F$ .

ITEMooQKISooBpyyee

(3) Si  $(x_n)$  est une suite bornée dans  $E$ , alors nous pouvons en extraire une sous-suite  $(x_{\varphi(n)})$  telle que  $y_n = f(x_{\varphi(n)})$  converge dans  $F$ .

Une application vérifiant les conditions équivalentes de la proposition 27.3 est dite **compacte**.

*Démonstration.* En trois parties.

(1) **(1)  $\Rightarrow$  (2)** La boule unité fermée est bornée, et donc son image est relativement compacte.

(2) **(2)  $\Rightarrow$  (3)** Soient  $(x_n)$  bornée dans  $E$ , et soit  $M$ , une borne. Nous avons donc

$$z_n = f\left(\frac{x_n}{M}\right) \in f(B(0, 1)) \subset \overline{f(B(0, 1))}. \quad (27.1)$$

Par hypothèse, la partie  $\overline{f(B(0, 1))}$  est compacte ; la suite  $(z_n)$  est contenue dans ce compact. Nous en déduisons que  $(z_n)$  admet une sous-suite convergente. Soit donc  $\varphi: \mathbb{N} \rightarrow \mathbb{N}$  telle que

$$f\left(\frac{x_{\varphi(n)}}{M}\right) \rightarrow y. \quad (27.2)$$

1. Définition 7.51.

2. Relativement compact, définition 7.129.

Vu que  $f$  est linéaire, nous avons aussi  $f(x_{\varphi(n)}) \rightarrow My$ .

- (3) **(3)  $\Rightarrow$  (1)** Soit une partie bornée  $A$  de  $F$ . Nous prouvons que  $\overline{f(A)}$  est compact. Pour cela nous considérons une suite  $(y_n)$  dans  $\overline{f(A)}$ . Soit  $\epsilon_k \rightarrow 0$ . Pour chaque  $n$ , l'élément  $y_n$  est dans la fermeture de  $f(A)$ , c'est-à-dire qu'il existe un élément de  $f(A)$  dont la distance à  $y_n$  est plus petite que  $\epsilon_n$ . Bref, il existe une suite  $(x_n)$  dans  $A$  telle que

$$\|f(x_n) - y_n\| \leq \epsilon_k. \quad (27.3)$$

La suite  $(x_n)$  est bornée parce qu'elle est contenue dans  $A$ . Donc il existe une sous-suite  $\varphi: \mathbb{N} \rightarrow \mathbb{N}$  telle que

$$f(x_{\varphi(n)}) \rightarrow y \quad (27.4)$$

pour un certain  $y \in \overline{f(A)}$ . Le fait que  $y$  soit dans  $\overline{f(A)}$  est dû au fait que la suite  $n \mapsto f(x_n)$  est une suite dans  $f(A)$ ; donc la limite est dans la fermeture de  $f(A)$ .

Nous avons

$$\|y_{\varphi(n)} - y\| \leq \|y_{\varphi(n)} - f(x_{\varphi(n)})\| + \|f(x_{\varphi(n)}) - y\| \quad (27.5a)$$

$$\leq \epsilon_k + \|f(x_{\varphi(n)}) - y\| \rightarrow 0. \quad (27.5b)$$

Donc la suite  $(y_{\varphi(n)})$  converge vers  $y \in \overline{f(A)}$ . Nous avons montré que toute dans  $\overline{f(A)}$  possédait une sous-suite convergente dans  $\overline{f(A)}$ . Le théorème de Bolzano-Weierstrass 7.194 conclut que  $\overline{f(A)}$  est compact et donc que  $f(A)$  est relativement compact.

□

LEMooQCJTTooNmMXEC

#### Lemme 27.4 ([1]).

Soient un espace métrique  $(E, d)$  ainsi qu'un espace compact  $K$ . Nous considérons la distance uniforme  $d_\infty$  sur  $C^0(K, E)$ . L'application d'évaluation

$$\begin{aligned} \varphi: (C^0(K, E), d_\infty) &\rightarrow (E, d) \\ f &\mapsto f(x). \end{aligned} \quad (27.6)$$

est continue.

*Démonstration.* Soit  $y \in E$ . Tout ouvert contenant  $y$  contient une boule  $B(y, r)$ . Nous montrons que  $\varphi^{-1}(B(y, r))$  est ouverte. Soit  $f_0 \in \varphi^{-1}(B(y, r))$ . Nous allons trouver un  $\epsilon > 0$  tel que  $B(f_0, \epsilon) \subset \varphi^{-1}(B(y, r))$ . Les éléments  $f_0$ ,  $x$  et  $y$  sont fixés, et nous avons  $d(f_0(x), y) < r$ . Il existe donc  $\alpha > 0$  tel que<sup>3</sup>

$$d(f_0(x), y) < \alpha < r. \quad (27.7)$$

Nous choisissons  $\epsilon < r - \alpha$  et nous calculons un peu :

$$d(g(x), y) \leq d(g(x), f_0(x)) + d(f_0(x), y) \quad (27.8a)$$

$$= \epsilon + d(f_0(x), y) \quad (27.8b)$$

$$< \epsilon + \alpha \quad (27.8c)$$

$$< r. \quad (27.8d)$$

Donc  $B(f_0, \epsilon) \subset \varphi^{-1}(B(y, r))$ .

Cela étant fait, nous prouvons que  $\varphi$  est continue. Soit un ouvert  $A$  de  $E$ . Pour montrer que  $\varphi^{-1}(A)$  est ouvert dans  $C^0(K, E)$ , nous considérons  $f_0 \in \varphi^{-1}(A)$  et nous allons trouver un ouvert  $S$  tel que  $f_0 \in S \subset \varphi^{-1}(A)$ .

Nous avons  $f_0(x) \in A$ ; nous considérons une boule  $B(f_0(x), r) \subset A$ , et nous allons trouver un ouvert  $S \subset C^0(K, E)$  tel que  $f_0 \in S \subset \varphi^{-1}(A)$ . La partie  $A$  étant ouverte dans le métrique  $E$  nous pouvons considérer  $r > 0$  tel que  $B(f_0(x), r) \subset A$ . □

3. Par exemple par le lemme 1.544(1).

LEMooKWGOooKzqDud

**Lemme 27.5** ([1]).

Soient des espaces topologiques  $X$  et  $Y$ . Nous considérons la topologie de la norme uniforme sur  $C^0(X, Y)$ . Soit une partie  $A$  de  $C^0(X, Y)$ . Si  $f \in \text{Adh}_{C^0(X, Y)}(A)$ , alors  $f(x) \in \text{Adh}_Y(A(x))$ .

*Démonstration.* Si  $f \in \text{Adh}(A)$ , alors il existe une suite  $(f_n)$  dans  $A$  telle que  $f_n \xrightarrow{C^0(X, Y)} f$ , et donc telle que  $f_n(x) \xrightarrow{Y} f(x)$ . Vu que chacun des  $f_n(x)$  est dans  $A(x)$ , nous avons bien  $f(x) \in \text{Adh}_Y(A(x))$ .  $\square$

PROPooCGBHooCpKycz

**Proposition 27.6** ([290]).

Soient des espaces topologiques  $X$  et  $Y$ . Nous considérons une suite équicontinue d'applications  $f_n: X \rightarrow Y$  telle que  $f_n \xrightarrow{\text{simple}} f$ . Alors la convergence est uniforme sur tout compact.

ThoKRbtph

**Théorème 27.7** (Théorème d'Ascoli[716]).

Soient un espace topologique compact  $K$  et un espace métrique  $(E, d)$ . Nous considérons la topologie de la distance uniforme<sup>4</sup> sur  $C^0(K, E)$ . Une partie  $A$  de  $C^0(K, E)$  est relativement compacte<sup>5</sup> si et seulement si les deux conditions suivantes sont remplies :

- (1)  $A$  est équicontinu<sup>6</sup>,
- (2)  $\forall x \in K$ , l'ensemble  $A(x) = \{f(x) \text{ tel que } f \in A\}$  est relativement compact dans  $E$ .

*Démonstration.* Attention, ça va être un peu long.

- (1)  $\Rightarrow$  Nous supposons que  $A \subset C^0(K, E)$  est relativement compact. Nous notons  $B$  un compact contenant  $A$ . Soit  $x \in K$ . L'application d'évaluation

$$\begin{aligned} \varphi: (C^0(K, E), d_\infty) &\rightarrow (E, d) \\ f &\mapsto f(x). \end{aligned} \quad (27.9)$$

est continue par le lemme 27.4. La partie  $B$  est compacte dans  $C^0(K, E)$ . Donc  $B(x) = \varphi(B)$  est compacte dans  $E$  par le théorème 7.274. Vu que  $A(x) \subset B(x)$  nous avons montré que  $A(x)$  est relativement compact.

Nous devons montrer que  $A$  est équicontinu. Nous prouvons que  $B$  est équicontinu en  $x$  en utilisant la caractérisation du lemme 7.372. Vu que  $B$  est compact, il existe un nombre fini de  $f_0, \dots, f_n \in B$  tels que

$$B \subset \bigcup_{k=0}^n B(f_k, \epsilon) \quad (27.10)$$

où  $B(f_i, \epsilon)$  est la boule dans  $(C^0(K, E), d_\infty)$ . Chacune des  $f_i$  est continue en  $x$ ; donc il existe un voisinage  $V_i$  de  $x$  tel que  $d(f_i(x), f_i(y)) < \epsilon$  pour tout  $y \in V_i$ . Nous posons

$$V = \bigcap_{i=1}^n V_i \quad (27.11)$$

qui, en tant que réunion finie d'ouverts contenant  $x$ , est un ouvert contenant  $x$ . Pour  $f \in B$  et  $y \in V$  nous avons

$$d(f(x), f(y)) \leq d(f(x), f_j(x)) + d(f_j(x), f_j(y)) + d(f_j(y), f(y)) < 3\epsilon. \quad (27.12)$$

- (2)  $\Leftarrow$  En plusieurs parties. Nous considérons l'adhérence  $B$  de  $A$  dans  $C^0(K, E)$ .

4. Définition 7.275.

5. Sa fermeture est compacte, définition 7.129.

6. Définition 7.370.

- (2a) **A est précompact** Nous montrons que  $A$  est précompact (définition 7.147). Soit  $\epsilon > 0$ . Étant donné que  $A$  est équicontinu, pour chaque  $x \in K$ , il existe un voisinage ouvert  $U_x$  de  $x$  tel que

$$d(f(y), f(x)) < \epsilon \quad (27.13)$$

pour tout  $y \in U_x$  et pour tout  $f \in A$ . Les parties  $\{U_x\}_{x \in K}$  forment un recouvrement du compact  $K$  par des ouverts. Il existe donc  $x_1, \dots, x_n \in K$  tels que  $U_{x_1}, \dots, U_{x_n}$  recouvrent  $K$ .

Nous posons

$$L = A(x_1) \cup \dots \cup A(x_n). \quad (27.14)$$

La partie  $L$  est relativement compacte parce qu'elle est contenue dans  $B_{x_1} \cup \dots \cup B_{x_n}$  qui est relativement compact comme union de relativement compacts.

- (2a.i) **Les parties  $C_a$**

Par relative compacité, il existe une partie finie  $J$  de  $E$  telle que

$$L \subset \bigcup_{s \in J} B(s, \epsilon). \quad (27.15)$$

Pour  $a = (a_1, \dots, a_n) \in J^n$  nous posons

$$C_a = \{f \in C^0(K, E) \text{ tel que } \forall i = 1, \dots, n, \forall y \in U_{x_i}, d(f(y), a_i) < 2\epsilon\}. \quad (27.16)$$

Notons que  $J^n$  étant fini, il existe seulement une quantité finie de telles parties.

- (2a.ii) **Les  $C_a$  recouvrent  $A$**  Nous vérifions maintenant que les  $\{C_a\}_{a \in J^n}$  recouvrent  $A$ . Soit  $f \in A$ . Vu que  $f(x_i) \in A(x_i) \subset L$  nous avons  $f(x_i) \in B(s_i, \epsilon)$  pour un certain  $s_i \in J \subset E$ . Nous considérons l'élément  $s = (s_1, \dots, s_n) \in J^n$ . Nous prouvons que  $f \in C_s$ . En effet soit  $i = 1, \dots, n$  et  $y \in U_{x_i}$ . Nous avons

$$d(f(y), s_i) \leq d(f(y), f(x_i)) + d(f(x_i), s_i). \quad (27.17)$$

D'abord  $d(f(x_i), s_i) \leq \epsilon$  parce que  $s_i$  a été choisit pour que  $f(x_i) \in B(s_i, \epsilon)$ . Ensuite  $d(f(y), f(x_i)) \leq \epsilon$  parce que  $y \in U_{x_i}$ . Bref,  $d(f(y), s_i) \leq 2\epsilon$  et  $f \in C_s$ .

- (2a.iii) **A est précompact** Chacun des  $C_a$  est contenu dans une boule de taille  $2\epsilon$ . Donc  $A$  est recouvert par un nombre fini de boules de rayon  $2\epsilon$ .

- (2b) **B est précompact** Pour rappel,  $B$  est l'adhérence de  $A$  dans  $C^0(K, E)$ . La proposition 7.148 dit que  $B$  est précompacte en tant qu'adhérence du précompact  $A$ .

- (2c) **B est complet** Soit une suite de Cauchy  $(f_n)$  dans  $\text{Adh}_{C^0(K, E)}(A)$ . Alors, étant donné le lemme 27.5, pour chaque  $x \in K$ , la suite  $(f_n(x))$  est une suite dans  $\text{Adh}_E(A(x))$ . Montrons qu'elle est même une suite de Cauchy. Soit  $\epsilon > 0$ . Il existe  $N \in \mathbb{N}$  tel que si  $m, n \geq N$ , alors  $d_\infty(f_m, f_n) < \epsilon$ . Pour de tels  $N, m, n$  nous avons

$$d(f_n(x), f_m(x)) \leq \sup_{y \in K} d(f_n(y), f_m(y)) = d_\infty(f_m, f_n) < \epsilon. \quad (27.18)$$

La partie  $A(x)$  étant relativement compacte, elle est contenue dans un compact  $S$  de  $E$ . Il se fait que  $S$  est complet par la proposition 7.314. Donc  $f_n(x)$  converge dans  $E$ . Nous notons  $f(x)$  cette limite. Cela nous définit une application  $f: K \rightarrow E$ .

Nous avons  $f_n \xrightarrow{\text{simple}} f$ , mais n'indique pour l'instant que  $f \in C^0(K, X)$  et encore moins que  $f_n \xrightarrow{C^0(K, X)} f$ . C'est ce que nous voyons maintenant.

D'après la proposition 27.6, la convergence est uniforme sur tout compact. Or justement  $f_n \rightarrow f$  est une convergence simple sur le compact  $K$ . Donc c'est une convergence uniforme et la limite est continue par la proposition 12.382. La convergence uniforme étant également celle de  $C^0(K, E)$ , nous avons bien  $f_n \xrightarrow{C^0(K, E)} f$ . Et donc  $f \in \text{Adh}_{C^0(K, E)}(A)$  parce que  $f_n \in \text{Adh}_{C^0(K, E)}(A)$  pour chaque  $n$ .

Cela montre que  $B$  est complet.

- (2d)  **$B$  est compact** Nous savons que  $B$  est précompact et complet. La proposition 7.149 dit qu'alors  $B$  est compact.
- (2e) **Conclusion** Vu que l'adhérence de  $A$  est compacte, nous déduisons que  $A$  est précompact.

□

La version suivante du théorème de Banach-Steinhaus est énoncée de façon *ad hoc* pour fonctionner avec l'espace  $\mathcal{D}(K)$  des fonctions de classe  $C^\infty$  à support dans le compact  $K$ . Un énoncé un peu plus fort est donné dans le cadre des espaces de Fréchet dans [259].

ThoNBrmGIg

**Théorème 27.8** (Banach-Steinhaus avec des seminormes[1]).

Soit  $(E, d)$  un espace vectoriel métrique complet dont la topologie est également<sup>7</sup> donnée par une famille  $\mathcal{P} = \{p_i\}_{i \in I}$  de seminormes. Soit  $\{T_\alpha\}_{\alpha \in A}$  une famille d'applications linéaires continues  $T_\alpha: E \rightarrow \mathbb{R}$  telles que pour tout  $x \in E$  nous ayons

$$\sup_{\alpha \in A} |T_\alpha(x)| < \infty. \quad (27.19)$$

Alors il existe une constante  $K > 0$  et un sous-ensemble fini  $J \subset I$  tels que pour tout  $x \in E$  et pour tout  $\alpha \in A$  nous ayons

$$|T_\alpha(x)| \leq K \max_{j \in J} p_j(x). \quad (27.20)$$

EqIFNGbtr

*Démonstration.* Pour chaque  $k \in \mathbb{N} \setminus \{0\}$  nous posons

$$\Omega_k = \{x \in E \text{ tel que } \sup_{\alpha \in A} |T_\alpha(x)| > k\}. \quad (27.21)$$

Ces ensembles sont des ouverts (pour la même raison que dans la preuve du théorème 10.175) et leur union est  $E$  en entier parce que par hypothèse  $\sup_{\alpha \in A} |T_\alpha(x)| < \infty$ .

- (1) **Les  $\Omega_k$  ne sont pas tous denses** Le théorème 7.419 nous dit que  $(E, d)$  est un espace de Baire. Supposons que les  $\Omega_k$  sont tous denses. Dans ce cas, l'intersection des  $\Omega_k$  est également dense (c'est la définition 7.417 d'un espace de Baire), et en particulier non vide.

Soit  $x_0 \in \bigcap_{k \in \mathbb{N}} \Omega_k$ . Nous avons

$$\sup_{\alpha \in A} |T_\alpha(x_0)| = \infty, \quad (27.22)$$

ce qui contredirait l'hypothèse. Donc les  $\Omega_k$  ne sont pas tous denses.

- (2) **Un non dense** Soit  $k_0 \in \mathbb{N}^*$  tel que  $\Omega_{k_0}$  n'est pas dense dans  $E$ . Il existe donc  $x_0 \in E$  et un ouvert  $\mathcal{O}$  autour de  $x_0$  n'intersectant pas  $\Omega_{k_0}$ .

- (3) **Topologie**

Parlons de topologie sur  $E$ . L'ouvert  $\mathcal{O}$  est un  $d$ -ouvert et donc un  $\mathcal{P}$ -ouvert, lequel contient une  $\mathcal{P}$ -boule ouverte<sup>8</sup>. Cette dernière boule n'est pas spécialement une  $d$ -boule, mais c'est un  $d$ -ouvert.

- (4) **Une majoration**

Il existe donc  $J$  fini dans  $I$  et  $\rho > 0$  tels que  $B_J(x_0, \rho) \subset \mathcal{O}$  et donc tels que

$$B_J(x_0, \rho) \cap \Omega_{k_0} = \emptyset. \quad (27.23)$$

Donc pour tout  $z \in B_J(x_0, \rho)$  nous avons

$$\sup_{\alpha \in A} |T_\alpha(z)| \leq k_0. \quad (27.24)$$

7. Au sens où les ouverts sont les mêmes.

8. Définition 7.396.

Si maintenant  $y \in B_J(0, \rho)$ , nous avons  $y = (x_0 + y) - x_0$  et donc

$$|T_\alpha(y)| = |T_\alpha(x_0 + y) - T_\alpha(x_0)| \quad (27.25a)$$

$$\leq |T_\alpha(x_0 + y)| + |T_\alpha(x_0)| \quad (27.25b)$$

$$\leq k_0 + C. \quad (27.25c)$$

Justifications.

—  $x_0 + y \in B_J(x_0, \rho)$ , donc  $|T_\alpha(x_0 + y)| \leq k_0$ .

— nous avons posé  $C = \sup_{\alpha \in A} |T_\alpha(x_0)|$

(5) **Sur la boule unité** En divisant par  $\rho$ , pour  $y \in B_J(0, 1)$  nous avons

$$|T_\alpha(y)| \leq \rho^{-1}(k_0 + C). \quad (27.26)$$

(6) **Un bon élément dans la boule unité** Pour n'importe quel  $x \in E$  nous avons

$$\frac{x}{\max_{j \in J} p_j(x)} \in B_J(0, 1). \quad (27.27)$$

Nous pouvons donc appliquer (27.26) à cet élément :

$$\left| T_\alpha \left( \frac{x}{\max_{j \in J} p_j(x)} \right) \right| \leq \rho^{-1}(k_0 + C). \quad (27.28)$$

Utilisant encore la linéarité de  $T_\alpha$  nous trouvons ce que nous devons trouver :

$$|T_\alpha(x)| \leq \max_{j \in J} p_j(x) \rho^{-1}(k_0 + C) = K \max_{j \in J} p_j(x) \quad (27.29)$$

en posant  $K = \rho^{-1}(k_0 + C)$ .

□

CorPGwLluz

**Corolaire 27.9** ([1, 220]).

Soit  $(X, \{p_i\}_{i \in I})$ , un espace vectoriel sur  $\mathbb{C}$  seminormé qui est

(1) localement compact<sup>9</sup>,

(2) métrisable et complet.

Soit  $(T_j)_{j \in \mathbb{N}}$  une suite d'application linéaires continues  $X \rightarrow \mathbb{C}$  telles que pour tout  $x \in X$  il existe<sup>10</sup>  $\alpha_x \in \mathbb{C}$  tel que

$$T_j x \xrightarrow{\mathbb{C}} \alpha_x. \quad (27.30)$$

Si nous posons  $Tx = \alpha_x$  alors

(1) L'application  $T$  est linéaire et continue.

(2) Si  $x_k \xrightarrow{X} x$ , alors

$$|T_k(x_k) - T(x_k)| \xrightarrow{\mathbb{R}} 0. \quad (27.31)$$

(3) Pour tout compact  $K$  dans  $X$  nous avons

$$\sup_{x \in K} |T_k x - Tx| \xrightarrow{\mathbb{R}} 0. \quad (27.32)$$

(4) si  $x_k \rightarrow x$  dans  $X$  alors

$$T_k x_k \xrightarrow{\mathbb{C}} Tx. \quad (27.33)$$

*Démonstration.* En plusieurs parties.

9. Définition 7.181.

10. L'unicité des  $\alpha_x$  et donc le fait que tout le reste ait un sens provient de l'hypothèse de séparabilité et la proposition 7.82.

- (1) **Linéaire** Vu que  $X$  est seminormé et que  $\mathbb{C}$  est un corps valué, la proposition 7.399 nous indique que  $X$  est un espace vectoriel topologique et la proposition 7.259 nous dit que la limite de suite est une opération linéaire. Pour  $x, y \in X$  nous pouvons donc faire

$$T(x + y) = \lim_{j \rightarrow \infty} (T_j(x + y)) \quad (27.34a)$$

$$= \lim_{j \rightarrow \infty} (T_j(x) + T_j(y)) \quad (27.34b)$$

$$= \lim_{j \rightarrow \infty} T_j(x) + \lim_{j \rightarrow \infty} T_j(y) \quad (27.34c)$$

$$= T(x) + T(y). \quad (27.34d)$$

De même nous avons

$$T(\lambda x) = \lim_{k \rightarrow \infty} (T_k(\lambda x)) \quad (27.35a)$$

$$= \lim_{k \rightarrow \infty} (\lambda T_k(x)) \quad (27.35b)$$

$$= \lambda \lim_{k \rightarrow \infty} T_k(x) \quad (27.35c)$$

$$= \lambda T(x). \quad (27.35d)$$

- (2) **Séquentiellement continue** Nous prouvons que  $T$  est séquentiellement continue<sup>11</sup>. Soit une suite  $x_k \xrightarrow{X} x$ . Nous voulons utiliser le théorème de Banach-Steinhaus 27.8. L'ensemble  $\{T_k\}_{k \in \mathbb{N}}$  est un ensemble d'applications linéaires continues. De plus, si  $x \in X$ , la suite  $k \mapsto T_k(x)$  est une suite convergente dans  $\mathbb{C}$ ; elle est donc bornée :

$$\sup_{k \in \mathbb{N}} |T_k(x)| < \infty. \quad (27.36)$$

Le théorème de Banach-Steinhaus nous dit qu'il existe un  $C > 0$  et  $J$  fini dans  $I$  tels que pour tout  $k$  et pour tout  $x$  nous ayons

$$|T_k(x)| \leq C \max_{j \in J} p_j(x). \quad (27.37)$$

Un seul  $i \in J$  nous suffira. Soient donc  $C > 0$  et  $i \in I$  tels que  $|T_k(x)| \leq Cp_i(x)$  pour tout  $x$ . Vu que le module est une opération continue sur  $\mathbb{C}$ , elle commute avec la limite et nous pouvons faire le calcul suivant :

$$|T(x)| = \left| \lim_{k \rightarrow \infty} T_k(x) \right| \quad (27.38a)$$

$$= \lim_{k \rightarrow \infty} |T_k(x)| \quad (27.38b)$$

$$\leq \lim_{k \rightarrow \infty} Cp_i(x) \quad (27.38c)$$

$$= Cp_i(x). \quad (27.38d)$$

Avec ça nous pouvons prouver que  $T$  est séquentiellement continue. Soit  $x_k \xrightarrow{X} 0$ . En utilisant la proposition 7.398 nous avons :

$$|T(x_k)| \leq Cp_i(x_k) \xrightarrow{\mathbb{R}} 0 \quad (27.39)$$

Donc  $|T(x_k)| \rightarrow 0$  et  $T$  est séquentiellement continue en zéro. Vu que  $T$  est linéaire, elle est séquentiellement continue partout.

- (3) **Continue** Vu que l'espace  $X$  est métrisable, la proposition 7.157 conclut que  $T$  est continue.

---

11. Définition 7.150



(4) **Point (2)**

Nous passons à la preuve de (2). Supposons que  $x_k \xrightarrow{X} x$ . Voici des majorations avec justifications juste en-dessous :

$$|T_k(x_k) - T(x_k)| \leq |T_k(x_k) - T_k(x)| + |T_k(x) - T(x_k)| \quad (27.40a)$$

$$= |T_k(x_k - x)| + |T_k(x) - T(x_k)| \quad (27.40b)$$

$$\leq Cp_i(x_k - x) + |T_k(x) - T(x_k)| \quad (27.40c)$$

$$\leq C\epsilon + |T_k(x) - T(x_k)| \quad (27.40d)$$

$$\leq C\epsilon + |T_k(x) - T(x)| + |T(x) - T(x_k)| \quad (27.40e)$$

$$\leq (C + 2)\epsilon. \quad (27.40f)$$

Justifications :

- Pour (27.40c). C'est le théorème de Banach-Steinhaus 27.8 qui nous assure l'existence d'un  $i \in I$  et d'une  $C \in \mathbb{R}^+$  qui donne la majoration.
- Pour (27.40d). Nous nous sommes fixé un  $\epsilon > 0$  et nous avons pris  $k$  assez grand. Pour la topologie des seminormes, la convergence  $y_k \xrightarrow{X} y$  implique la limite  $p_i(y_k - y) \xrightarrow{\mathbb{R}} 0$  pour toutes les seminormes, et en particulier pour celle donnée par le théorème de Banach-Steinhaus.
- Pour (27.40f). D'une part, par hypothèse et par définition de  $T$ , nous avons  $T_k(x) \xrightarrow{\mathbb{C}} T(x)$  et d'autre part par continuité séquentielle de  $T$  et par convergence de  $x_k \xrightarrow{X} x$  nous avons  $T(x_k) \xrightarrow{\mathbb{C}} T(x)$ .

En prenant  $k$  encore plus grand, nous avons toutes les majorations en même temps.

Nous avons donc bien

$$|T_k(x_k) - T(x_k)| \xrightarrow{\mathbb{R}} 0. \quad (27.41)$$

(5) **Point (3)** Nous nous lançons dans la preuve du point (3). En vertu du point (1), l'application

$$\begin{aligned} S_k: K &\rightarrow \mathbb{R} \\ x &\mapsto |T_k(x) - T(x)| \end{aligned} \quad (27.42)$$

est continue. Vu que  $K$  est un compact, le théorème 7.196, nous indique qu'il existe  $x_k$  tel que

$$|T_k(x_k) - T(x_k)| = \sup_{x \in K} |T_k(x) - T(x)|. \quad (27.43)$$

Nous considérons la suite

$$a_k = |T_k(x_k) - T(x_k)| \quad (27.44)$$

et nous allons montrer que

- toutes les sous-suites convergentes de  $(a_k)$  convergent vers 0.
- la suite  $(a_k)$  est contenue dans un compact de  $\mathbb{R}$ .

Ensuite le corolaire 7.343 terminera.

(5a) **Toutes les sous-suites convergentes** Soit  $\varphi: \mathbb{N} \rightarrow \mathbb{N}$  telle que la suite  $b_k = a_{\varphi(k)}$  soit convergente. Nous allons trouver une sous-suite qui converge vers 0. La suite  $k \mapsto a_{\varphi(k)}$  est une suite dans le compact  $K$  ; elle contient donc une sous-suite convergente. Soit donc  $\psi: \mathbb{N} \rightarrow \mathbb{N}$  telle que  $k \mapsto a_{(\varphi \circ \psi)(k)}$  soit convergente. Nous étudions la sous-suite  $c_k = b_{\psi(k)}$  et nous allons prouver que  $c_k \rightarrow 0$ .

Pour ne pas mourir sous les notations, nous posons  $\sigma = \varphi \circ \psi$ . Nous avons donc  $x_{\sigma(k)} \xrightarrow{X} x$ , et donc  $|T_{\sigma(k)}(x_{\sigma(k)}) - T(x_{\sigma(k)})| \xrightarrow{\mathbb{R}} 0$  par le point (2).

Donc la suite  $a_{\varphi(k)}$  qui est convergente contient une sous-suite qui converge vers 0. La suite  $a_{\varphi(k)}$  converge donc vers 0.

- (5b) **Dans un compact** Nous allons prouver que  $|T_k(x_k) - T(x_k)|$  est bornée. Supposons que ce ne soit pas le cas. Pour chaque  $n \in \mathbb{N}$ , il existe un  $k$  tel que  $|T_k(x_k) - T(x_k)| > n$ . En numérotant bien, il existe une application  $\varphi: \mathbb{N} \rightarrow \mathbb{N}$  telle que

$$|T_{\varphi(k)}(x_{\varphi(k)}) - T(x_{\varphi(k)})| > k. \quad (27.45)$$

La suite  $k \mapsto x_{\varphi(k)}$  est une suite dans le compact  $K$ . Elle admet une sous-suite convergente. En nommant  $x_{\sigma(k)}$  cette sous-suite, nous avons  $x_{\sigma(k)} \xrightarrow{X} x$ , et nous nous avons encore  $|T_{\sigma(k)}(x_{\sigma(k)}) - T(x_{\sigma(k)})| \xrightarrow{\mathbb{R}} 0$ .

Ouais mais bon. Une suite qui vérifie (27.45) ne peut pas avoir une sous-suite convergente. Contradiction.

- (6) **Point (4)** Nous avons la majoration

$$|T_k x_k - T x| \leq |T_k x_k - T x_k| + |T x_k - T x|. \quad (27.46)$$

Par le point (2), le premier terme tend vers zéro. La continuité de  $T$  (point (1)) s'occupe du deuxième terme.

□

## 27.3 Espaces de Lebesgue $L^p$

### 27.3.1 Généralités

#### Définition 27.10.

Soit  $(\Omega, \mathcal{F}, \mu)$  un espace mesuré. Deux fonctions à valeurs complexes  $f$  et  $g$  sur cet espace sont dites **équivalentes** et nous notons  $f \sim g$  si elles sont  $\mu$ -presque partout égales. Nous notons  $[f]$  la classe de  $f$  pour cette relation.

Nous allons souvent noter  $f$  indifféremment pour la fonction  $f$  et un des représentants de la classe de  $f$ . Toutefois, lorsque la distinction sera importante, nous essayerons de faire faire l'effort de distinguer la fonction  $f$  de sa classe  $[f]$ .

#### Lemme 27.11.

Une classe contient au maximum une seule fonction continue.

*Démonstration.* Soient deux fonctions continues  $f_1$  et  $f_2$  avec  $f_1(a) \neq f_2(a)$ . Si  $|f_1(a) - f_2(a)| > \delta$  alors il existe un  $\epsilon$  tel que  $|f_1(x) - f_1(a)| > \delta$  pour tout  $x \in B(a, \epsilon)$ . En particulier  $f_1 \neq f_2$  sur  $B(a, \epsilon)$ . Cette dernière boule est de mesure de Lebesgue non nulle ; ergo  $f_1$  et  $f_2$  ne sont pas dans la même classe.

□

#### Définition 27.12.

Pour  $p \geq 1$ , nous introduisons l'opération

$$\|f\|_p = \left( \int_{\Omega} |f(x)|^p d\mu(x) \right)^{1/p} \quad (27.47)$$

et nous notons  $\mathcal{L}^p(\Omega, \mu)$  l'ensemble des fonctions mesurables sur  $\Omega$  telles que  $\|f\|_p < \infty$ .

#### 27.13.

Le fait que  $f$  doive être mesurable pour être dans  $L^p$  est bien dans la définition de  $L^p$ , et non une propriété qui pourrait être déduite de la finitude de l'intégrale (27.47).

Soit  $\Omega = \mathbb{R}$  muni de ses boréliens ou même de sa tribu de Lebesgue et de la mesure de Lebesgue. Si  $A$  est une partie bornée non mesurable (exemple 14.157), alors nous considérons un borné mesurable  $K$  contenant  $A$  et la fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in A \\ -1 & \text{si } x \in K \setminus A \\ 0 & \text{sinon.} \end{cases} \quad (27.48)$$

Nous avons alors  $|f| = \mathbb{1}_K$ . Vu que  $K$  est borné,  $\int_{\mathbb{R}} |f| = \mu(K)$  ne présente pas de problèmes. Donc  $f$  serait un élément de  $L^1(\mathbb{R})$  sans être mesurable.

Pour éviter cela, nous incluons la mesurabilité dans la définition de  $\mathcal{L}^p$ .

LEMooHVZGooRwHXMk

**Lemme 27.14** ([1]).

*L'ensemble  $\mathcal{L}^p$  est un espace vectoriel sur  $\mathbb{C}$ .*

*Démonstration.* Pour rappel, nous ne considérons les choses que pour  $p \geq 1$ . Le fait que si  $f \in L^p$ , alors  $\lambda f \in L^p$  est évident. Ce qui est moins immédiat, c'est le fait que  $f + g \in L^p$  lorsque  $f$  et  $g$  sont dans  $L^p$ . Cela découle d'une part du fait que la fonction  $\varphi: x \mapsto x^p$  est convexe sur les positifs (lemme 17.87), de telle sorte que

$$\varphi\left(\frac{|a| + |b|}{2}\right) \leq \frac{\varphi(|a|) + \varphi(|b|)}{2}, \quad (27.49)$$

ou encore

$$(|a| + |b|)^p \leq 2^{p-1}(|a|^p + |b|^p) \quad (27.50)$$

EqZESduFa

Et d'autre part, nous savons que pour  $z_1, z_2 \in \mathbb{C}$ ,  $|z_1 + z_2| \leq |z_1| + |z_2|$  (proposition 9.108(1)). Donc

EqQooKRMEooSLHUUC

$$\int_{\Omega} |f + g|^p d\mu = \int_{\Omega} |f(\omega) + g(\omega)|^p d\mu(\omega) \quad (27.51a)$$

$$\leq \int_{\Omega} (|f(\omega)| + |g(\omega)|)^p d\mu(\omega) \quad (27.51b)$$

$$\leq 2^{p-1} \int_{\Omega} (|f(\omega)|^p + |g(\omega)|^p) d\mu(\omega) \quad (27.51c)$$

$$= 2^{p-1} (\|f\|_p^p + \|g\|_p^p). \quad (27.51d)$$

Donc si  $f$  et  $g$  sont dans  $\mathcal{L}^p$ , alors  $f + g$  est dans  $\mathcal{L}^p$ .  $\square$

### 27.15.

Il est à noter que nous ne considérons que des valeurs  $p \geq 1$ , précisément parce que la fonction  $x \mapsto |x|^p$  n'est pas convexe lorsque  $p < 1$ .

Dans le même ordre d'idées, si  $p \geq 1$ , alors le  $q \in \mathbb{R}$  tel que

$$\frac{1}{p} + \frac{1}{q} = 1 \quad (27.52)$$

est également  $q \geq 1$ . Cela est important pour un certain nombres de théorèmes qui vont venir, en particulier l'inégalité de Hölder (27.102).

Si vous en voulez à propos de  $0 < p < 1$ , vous pouvez lire [717].

### 27.16.

L'opération  $f \mapsto \|f\|_p$  n'est pas une norme sur  $\mathcal{L}^p$  parce que pour  $f$  presque partout nulle, nous avons  $\|f\|_p = 0$ . Il y a donc des fonctions non nulles sur lesquelles  $\|\cdot\|_p$  s'annule.

### 27.17.

Soit un espace mesuré non complet <sup>12</sup>  $(\Omega, \mathcal{A}, \mu)$ . Il existe une partie  $N$  de mesure nulle et  $A \subset N$  non mesurable. Considérons

$$f(x) = \begin{cases} 2 & \text{si } x \in A \\ 1 & \text{si } x \in N \setminus A \\ 0 & \text{sinon.} \end{cases} \quad (27.53)$$

Cette fonction est non nulle exactement sur  $N$ . Donc  $f \sim 0$ . Mais  $f$  n'est pas mesurable parce que  $f^{-1}(2) = A$  n'est pas mesurable.

Il est donc possible d'être dans la classe d'équivalence d'une fonction mesurable sans être mesurable. Ceci est cependant un détail (presque) sans importance pour deux raisons.

<sup>12</sup>. Définition 14.68.

- La mesure de Lebesgue est complète par définition (définition 14.143) si nous considérons bien la tribu de Lebesgue et non seulement les boréliens.
- Dans le cas des espaces de Lebesgue  $L^p(\Omega, \mathcal{A}, \mu)$ , il s'agit d'un quotient de  $\mathcal{L}^p$  qui ne contient que des fonctions mesurables. Donc dans l'étude de  $L^p$ , tous les représentants sont mesurables, même si  $(\Omega, \mathcal{A}, \mu)$  n'est pas complet.

LemKZVHVAR

**Lemme 27.18.**

Si  $f \in \mathcal{L}^p(\Omega)$  et  $f \sim g$ , alors  $g \in \mathcal{L}^p(\Omega)$  et  $\|f\|_p = \|g\|_p$ .

*Démonstration.* Soit  $h(x) = |g(x)|^p - |f(x)|^p$ ; c'est une fonction par hypothèse presque partout nulle et donc intégrable sur  $\Omega$ ; son intégrale y vaut zéro. Nous avons

$$\int_{\Omega} |f(x)|^p d\mu(x) = \int_{\Omega} (|f(x)|^p + h(x)) d\mu(x) = \int_{\Omega} |g(x)|^p d\mu(x). \quad (27.54)$$

Cela prouve que la dernière intégrale existe et vaut la même chose que la première.  $\square$

Nous pouvons donc considérer la norme  $|\cdot|_p$  comme une norme sur l'ensemble des classes plutôt que sur l'ensemble des fonctions. Nous notons  $L^p$  l'ensemble des classes des fonctions de  $\mathcal{L}^p$ . Cet espace est muni de la norme

$$\|[f]\|_p = \|f\|_p, \quad (27.55)$$

formule qui ne dépend pas du représentant par le lemme 27.18.

PROPooTYCYooAKJWOX

**Proposition 27.19.**

Soient un espace mesuré  $(\Omega, \mathcal{A}, \mu)$ , et  $p \geq 1$ .

(1) L'ensemble des classes  $L^p(\Omega, \mathcal{A}, \mu)$  est un espace vectoriel.

ITEMooZNXPOoGzFnIK

(2) La formule

$$\|[f]\|_p = \left( \int_{\Omega} |f(x)|^p d\mu(x) \right)^{1/p} \quad (27.56)$$

définit une norme<sup>13</sup> sur  $L^p(\Omega, \mathcal{A}, \mu)$ .

*Démonstration.* En deux parties.

(1) **Espace vectoriel** Le lemme 27.14 dit que  $\mathcal{L}^p$  est un espace vectoriel. Une structure d'espace vectoriel sur  $L^p(\Omega, \mathcal{A}, \mu)$  est donnée en posant

$$[f] + [g] = [f + g] \quad (27.57)$$

et

$$\lambda[f] = [\lambda f] \quad (27.58)$$

qui sont deux définitions correctes parce qu'elles ne dépendent pas du choix du représentant. De plus le lemme 27.14 dit que  $f + g \in \mathcal{L}^p$  dès que  $f, g \in \mathcal{L}^p$ . Donc  $[f + g] \in L^p$  dès que  $[f], [g] \in L^p$ .

(2) **Norme**

Pour être une norme, il faut vérifier les trois propriétés de la définition 7.212.

D'abord, si  $\|[f]\|_p = 0$ , nous avons

$$\int_{\Omega} |f(x)|^p d\mu(x) = 0, \quad (27.59)$$

ce qui par le lemme 14.199 implique que  $|f(x)|^p = 0$  pour presque tout  $x$ . Ou encore  $f \sim 0$ , c'est-à-dire  $[f] = [0]$  au niveau des classes.

13. Définition 7.212.

Ensuite pour  $\lambda \in \mathbb{C}$  et  $f \in L^p$  nous avons

$$\|\lambda f\|_p = \left( \int_{\Omega} |\lambda f(\omega)|^p d\mu(\omega) \right)^{1/p} \quad (27.60a)$$

$$= \left( \int_{\Omega} |\lambda|^p |f(\omega)|^p d\mu(\omega) \right)^{1/p} \quad (27.60b)$$

$$= |\lambda| \|f\|_p. \quad (27.60c)$$

Et enfin, en suivant le calcul (27.51) nous avons

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \quad (27.61)$$

□

## 27.20.

À partir de maintenant  $(L^p(\Omega, \mu), \|\cdot\|_p)$  est un espace métrique avec toute la topologie qui va avec.

Dans la suite nous n'allons pas toujours écrire  $[f]$  pour la classe de  $f$ . Par abus de notations nous allons souvent parler de  $f \in L^p$  comme si c'était une fonction.

De même nous notons  $L^p(\Omega)$  ou  $L^p(\Omega, \mu)$  ou  $L^p(\Omega, \mathcal{A}, \mu)$  d'après ce sur quoi nous voulons insister. Mais seule la dernière notation est parfaitement correcte.

PROP00RSDB00HEjQDq

### Proposition 27.21 ([1]).

Soit un espace mesuré  $\sigma$ -fini<sup>14</sup>  $(\Omega, \mathcal{A}, \mu)$ . Soit une fonction mesurable  $f: \Omega \rightarrow \mathbb{C}$ . Il existe des parties mesurables  $\Omega_n$  telles que pour tout  $n$

ITEM00XSIN00DDPzmb

$$(1) \mu(\Omega_n) < \infty,$$

ITEM00EGWE00AtEWfw

$$(2) \Omega_n \subset \Omega_{n+1}$$

ITEM00GEEZ00IjPQmv

$$(3) \int_{\Omega_n} |f| < \infty$$

ITEM00SDPT00OuiyOj

$$(4) \bigcup_n \Omega_n = \{|f| \neq \infty\}.$$

*Démonstration.* Vu que  $\Omega$  est  $\sigma$ -finie, il existe des parties mesurables  $S_n$  telles que  $\mu(S_n) < \infty$  et  $\bigcup_n S_n = \Omega$ . Nous commençons par poser  $A_n = \bigcup_{i=1}^n s_i$ . Les parties  $A_n$  vérifient déjà (1) et (2).

Pour les conditions (3) et (4), il suffit de poser  $\Omega_n = A_n \cap \{|f| < n\}$ . □

PROP00QNNV00TKRdyC

### Proposition 27.22 ([1]).

Soit un espace mesuré  $\sigma$ -fini  $(\Omega, \mathcal{A}, \mu)$ . Soit une fonction  $f: \Omega \rightarrow \mathbb{C}$  vérifiant  $\mu(|f| = \infty) = 0$ . Nous considérons des parties mesurables  $\Omega_k$  telles que

$$(1) \mu(\Omega_n) < \infty,$$

$$(2) \Omega_n \subset \Omega_{n+1}$$

$$(3) \int_{\Omega_n} |f| < \infty$$

$$(4) \bigcup_n \Omega_n = \{|f| \neq \infty\}.$$

Alors en notant  $\|f\|_{n,p}$  la norme  $L^p$  de  $f$  sur  $\Omega_n$ , nous avons

$$\|f\|_{k,p} \rightarrow \|f\|_p \quad (27.62)$$

*Démonstration.* Nous utilisons la convergence monotone 14.178 en posant  $g_k(x) = |f(x)|^p \mathbb{1}_{\Omega-k}$ . Les fonctions  $g_k$  sont à valeurs dans  $[0, \infty]$  et nous avons  $g_k \rightarrow |f|^p$  ponctuellement. Donc

$$\lim_{k \rightarrow \infty} \int_{\Omega_k} |f|^p = \lim_{k \rightarrow \infty} \int_{\Omega} g_k \quad (27.63a)$$

$$= \int_{\Omega} \lim_{k \rightarrow \infty} g_k \quad (27.63b)$$

$$= \int_{\Omega} |f|^p, \quad (27.63c)$$

ce qu'il fallait démontrer. □

---

14. Définition 14.11.

### 27.3.2 Un peu de convergence de suites

PropWoywYG

#### Proposition 27.23 ([718]).

Soit  $1 \leq p \leq \infty$  et supposons que la suite  $[f_n]$  dans  $L^p(\Omega, \mathcal{F}, \mu)$  converge vers  $[f]$  au sens  $L^p$ . Alors il existe une sous-suite  $(h_n)$  qui converge ponctuellement  $\mu$ -presque partout vers  $f$ .

*Démonstration.* Si  $p = \infty$  nous sommes en train de parler de la convergence uniforme et il ne faut même pas prendre ni de sous-suite ni de « presque partout ».

Supposons que  $1 \leq p < \infty$ . Nous considérons une sous-suite  $[h_n]$  de  $[f_n]$  telle que

$$\|[h_j] - [f]\|_p < 2^{-j}, \quad (27.64)$$

puis nous posons  $u_k(x) = |h_k(x) - f(x)|^p$ . Notons que ce  $u_k$  est une vraie fonction, pas une classe. Et en plus c'est une fonction positive. Nous avons

$$\int_{\Omega} u_k d\mu = \int_{\Omega} |h_k(x) - f(x)|^p d\mu(x) = \|h_k - f\|_p^p \leq 2^{-kp}. \quad (27.65)$$

Vu que  $u_k$  est une fonction positive la suite des sommes partielles de  $\sum_k u_k$  est croissante et vérifie donc le théorème de la convergence monotone 14.178 :

$$\int_{\Omega} \left( \sum_{k=0}^{\infty} u_k(x) \right) d\mu(x) = \sum_{k=0}^{\infty} \int_{\Omega} u_k(x) d\mu(x) \leq \sum_{k=0}^{\infty} 2^{-kp} < \infty. \quad (27.66)$$

Le fait que l'intégrale de la fonction  $\sum_k u_k$  est finie implique que cette fonction est finie  $\mu$ -presque partout. Donc le terme général tend vers zéro presque partout, c'est-à-dire

$$|h_k(x) - f(x)|^p \rightarrow 0. \quad (27.67)$$

Cela signifie que  $h_k \rightarrow f$  presque partout ponctuellement. □

Est-ce qu'on peut faire mieux que la convergence ponctuelle presque partout d'une sous-suite ? En tout cas on ne peut pas espérer grand chose comme convergence pour la suite elle-même, comme le montre l'exemple suivant.

ExPOmxICc

#### Exemple 27.24 ([719]).

Nous allons montrer une suite de fonctions qui converge vers zéro dans  $L^p[0, 1]$  (avec  $p < \infty$ ) mais qui ne converge ponctuellement pour *aucun* point.

Nous construisons la suite de fonctions par paquets. Le premier paquet est formé de la fonction constante 1.

Le second paquet est formé de deux fonctions. La première est  $\mathbb{1}_{[0, 1/2]}$  et la seconde  $\mathbb{1}_{[1/2, 1]}$ .

Plus généralement le paquet numéro  $k$  est constitué des  $k$  fonctions  $\mathbb{1}_{[i/k, (i+1)/k]}$  avec  $i = 0, \dots, k-1$ .

Vu que les fonctions du paquet numéro  $k$  ont pour norme  $\|f\|_p = \frac{1}{k}$ , nous avons évidemment  $f_n \rightarrow 0$  dans  $L^p$ . Il est par contre visible que chaque paquet passe en revue tous les points de  $[0, 1]$ . Donc pour tout  $x$  et pour tout  $N$ , il existe (même une infinité)  $n > N$  tel que  $f_n(x) = 1$ . Il n'y a donc convergence ponctuelle nulle part. △

La proposition suivante est une espèce de convergence dominée de Lebesgue pour  $L^p$ .

PropBVHXycL

#### Proposition 27.25.

Soit  $f \in L^p(\Omega)$  avec  $1 \leq p < \infty$  et  $(f_n)$  une suite de fonctions convergeant ponctuellement vers  $f$  et telle que  $|f_n| \leq |f|$ . Alors  $f_n \xrightarrow{L^p} f$ .

*Démonstration.* Nous avons immédiatement  $|f_n(x)|^p \leq |f(x)|^p$ , de telle sorte que le théorème de la convergence dominée implique que  $f_n \in L^p$ . La convergence dominée donne aussi que  $\|f_n\|_p \rightarrow \|f\|_p$ , mais cela ne nous intéresse pas ici.

Nous posons  $h_n(x) = |f_n(x) - f(x)|$ . En reprenant la formule de majoration (27.50) et en tenant compte du fait que  $|f_n(x)| \leq |f(x)|$ , nous avons

$$h_n(x) \leq 2^{p-1}(|f_n(x)|^p + |f(x)|^p) \leq 2^p |f(x)|^p, \quad (27.68)$$

ce qui prouve que  $|h_n|$  est uniformément (en  $n$ ) majorée par une fonction intégrable, donc  $h_n$  est intégrable et on peut permuter la limite et l'intégrale (théorème de la convergence dominée 14.207) :

$$\lim_{n \rightarrow \infty} \|f_n - f\|_p^p = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} |f_n(x) - f(x)|^p dx = \int_{\mathbb{R}^d} \lim_{n \rightarrow \infty} h_n(x) dx = 0. \quad (27.69)$$

□

PropRERZooYcEhc

### Proposition 27.26 ([1]).

Soit un espace mesuré  $(\Omega, \mathcal{A}, \mu)$  et une fonction mesurable  $f: \Omega \rightarrow \mathbb{C}$  telle que

$$\int_{\Omega} f \mathbb{1}_A d\mu = 0 \quad (27.70)$$

pour toute partie mesurable  $A$  de mesure finie. Alors  $[f] = 0$ , c'est-à-dire que  $f$  est non nulle uniquement sur une partie de mesure nulle.

*Démonstration.* Nous rappelons que dire que  $f$  est intégrable signifie que  $\operatorname{Re}(f)^+$ ,  $\operatorname{Re}(f)^-$ ,  $\operatorname{Im}(f)^+$  et  $\operatorname{Im}(f)^-$  sont intégrables.

Les parties

$$\{x \in \Omega \text{ tel que } \operatorname{Re}(f)^+(x) \in [k, k+1[ \}_{k=1, \dots} \quad (27.71)$$

et

$$\{x \in \Omega \text{ tel que } \operatorname{Re}(f)^+(x) \in [\frac{1}{k+1}, \frac{1}{k}[ \}_{k=1, \dots} \quad (27.72)$$

sont mesurables en tant qu'images inverses de mesurables par la fonction mesurable  $\operatorname{Re}(f)^+$ .

Nous notons  $\{A_i\}_{i \in \mathbb{N}}$  une énumération quelconque de ces parties. L'important est que

$$\bigcup_{i \in \mathbb{N}} A_i = \{x \in \Omega \text{ tel que } \operatorname{Re}(f)^+(x) \neq 0\} \quad (27.73)$$

et que pour chaque  $i$ , il existe  $\alpha_i > 0$  tel que  $\operatorname{Re}(f)^+(x) > \alpha_i$  pour tout  $x$  dans  $A_i$ .

Par hypothèse nous avons  $\int_{\Omega} f \mathbb{1}_{A_i} d\mu = 0$ . En particulier,

$$0 = \int_{\Omega} \operatorname{Re}(f \mathbb{1}_{A_i})^+ = \int_{\Omega} \operatorname{Re}(f)^+(x) \mathbb{1}_{A_i}(x) d\mu(x). \quad (27.74)$$

Mais pour tout  $x \in A_i$  nous avons  $\operatorname{Re}(f)^+(x) > \alpha_i > 0$ , donc

$$0 = \int_{\Omega} \operatorname{Re}(f)^+(x) \mathbb{1}_{A_i}(x) d\mu(x) \geq \alpha_i \int_{\Omega} \mathbb{1}_{A_i} = \alpha_i \mu(A_i) \geq 0. \quad (27.75)$$

Vu cet encadrement par zéro, nous avons  $\alpha_i \mu(A_i) = 0$  et donc  $\mu(A_i) = 0$  pour tout  $i$ .

Nous en déduisons, par union dénombrable de parties de mesures nulles, que  $\operatorname{Re}(f)^+$  est non nulle seulement sur une partie de mesure nulle.

Le même raisonnement pour  $\operatorname{Re}(f)^-$ ,  $\operatorname{Im}(f)^+$  et  $\operatorname{Im}(f)^-$  donne que  $f$  est non nulle sur une partie de mesure nulle. Donc  $f = 0$  au sens des classes dans  $L^p$ . □

### 27.3.3 L'espace $L^\infty$

SUBSECooYFJTtooBqrLXv

Ne pas confondre la norme suprénum (définition 10.107) avec la norme  $L^\infty$ , bien que les deux peuvent être écrites  $\|f\|_\infty$ .

**27.27.**

Il n'est pas possible de définir le supremum d'une fonction définie à ensemble de mesure nulle près parce que toute classe contient des fonctions qui peuvent être arbitrairement grandes en n'importe quel point. Nous cherchons alors à définir une notion de supremum qui ne tient pas compte des ensembles de mesure nulle.

DEFooIQOOooLpJBqi

**Définition 27.28.**

Soit  $f: \Omega \rightarrow \mathbb{C}$ . Un nombre  $M$  est un **majorant essentiel** de  $f$  si

$$\mu(|f(x)| \geq M) = 0. \quad (27.76)$$

Nous posons alors

$$N_\infty(f) = \inf\{M \text{ tel que } |f(x)| \leq M \text{ presque partout}\}. \quad (27.77)$$

Le nombre (peut-être infini)  $N_\infty(f)$  sera souvent écrite  $\|f\|_{L^\infty}$  ou  $\|f\|_\infty$ . Il est nommé le **suprémum essentiel** de  $f$ . Il pourra également être noté

$$\text{ess-sup}_{x \in \Omega} |f(x)|. \quad (27.78)$$

EQooZHKBooPgdzCp

Cela revient à prendre le supremum à ensemble de mesure nulle près.

DEFooXUKHooXYrlYq

**Définition 27.29.**

Nous définissons alors les espaces de Lebesgue correspondants :

$$\mathcal{L}^\infty(\Omega) = \{f: \Omega \rightarrow \mathbb{C} \text{ tel que } N_\infty(f) < \infty\}, \quad (27.79)$$

et  $L^\infty$  en est le quotient usuel.

**27.30.**

Le point sur les notations telles qu'elles devraient être respectées :

- (1) Si  $f$  est une fonction,  $N_\infty(f)$  est son supremum essentiel.
- (2) Si  $f$  est une fonction,  $\|f\|_\infty$  est sa norme supremum. Ce n'est pas la même chose que  $N_\infty(f)$ .
- (3) Si  $[f]$  est une classe de fonctions (pour l'égalité presque partout), alors  $\|[f]\|_{L^\infty}$  est la norme de cette classe dans  $L^\infty$ , c'est-à-dire  $\|[f]\|_{L^\infty} = N_\infty(f)$  où  $f$  est un représentant.

Le point sur les abus tolérables :

- (1) Si  $f$  est une fonction, on peut écrire  $\|f\|_{L^\infty}$  pour  $N_\infty(f)$ . Attention toutefois que  $N_\infty(f)$  peut valoir  $\infty$ , alors que les éléments de  $L^\infty$  sont sélectionnés pour être les classes des fonctions telles que  $N_\infty(f) < \infty$ . Donc il est parfois possible, lorsque  $f$  est une fonction, de parler de  $\|f\|_{L^\infty}$  alors que  $[f]$  n'est pas un élément de  $L^\infty$ .
- (2) Si  $[f]$  est une classe, on peut écrire  $N_\infty([f])$  pour  $\|[f]\|_{L^\infty}$ .
- (3) Noter  $f$  la classe de la fonction  $f$ . Attention qu'alors, écrire  $\|f\|_\infty$  n'a pas de sens.

Le point sur les abus intolérables :

- (1) Si  $f$  est une fonction, noter  $\|f\|_\infty$  pour  $N_\infty(f)$ .
- (2) Si  $[f]$  est une classe de fonctions, noter  $\|[f]\|_\infty$  pour  $\|[f]\|_{L^\infty}$ .

**27.31.**

Tout ceci pour dire que si  $f_k$  sont des fonctions et si  $f$  est une fonction, nous avons la convergence

$$f_k \xrightarrow{L^\infty} f \quad (27.80)$$

si et seulement si  $N_\infty(f_k - f) \rightarrow 0$ . Cette convergence (qui se sert des abus tolérables de notations) est équivalente à la convergence

$$[f_k] \xrightarrow{L^\infty} [f]. \quad (27.81)$$

Mais ce n'est pas du tout équivalent à  $\|f_k - f\|_\infty \rightarrow 0$ . Tout au plus, il est vrai que si  $\alpha \in L^\infty$  (donc  $\alpha$  est une classe), il existe un représentant  $f \in \alpha$  tel que  $\|f\|_\infty = \|\alpha\|_{L^\infty}$ .



**Lemme 27.32** ([1]).

Si  $f_k$  est une suite de fonctions telle que  $\|f_k - f\|_\infty \rightarrow 0$ , alors

$$f_k \xrightarrow{L^\infty} f \quad (27.82)$$

où la convergence signifie  $N_\infty(f_k - f) \rightarrow 0$ .

*Démonstration.* Si  $g$  est une fonction nous avons toujours  $N_\infty(g) \leq \|g\|_\infty$ . Donc

$$N_\infty(f_k - f) \leq \|f_k - f\|_\infty \rightarrow 0. \quad (27.83)$$

□

**27.33.**

Attention toutefois que ce lemme ne signifie pas que si  $\|f_k - f\|_\infty \rightarrow 0$ , alors  $[f_k] \xrightarrow{L^\infty} [f]$  parce que nous pourrions avoir  $N_\infty(f_k) = \infty$  et alors  $[f_k]$  n'est pas un élément de  $L^\infty$ .

Cela arrive par exemple pour  $f_k(x) = x$  et  $f(x) = x$ . Nous avons  $\|f_k - f\|_\infty = 0$  pour tout  $k$ , alors que ni  $f_k$  ni  $f$  ne donnent lieu à un élément de  $L^\infty$ .

**Exemple 27.34.**

Même si  $g$  est bornée, nous n'avons pas spécialement  $\|g\|_\infty = N_\infty(g)$ . Par exemple

$$\begin{aligned} g: \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{sinon} \end{cases} \end{aligned} \quad (27.84)$$

Cette fonction vérifie  $\|g\|_\infty = 1$  mais  $N_\infty(g) = 0$ .

△

**27.3.4 Quelques identifications**

Il est intuitivement clair que ce qui peut arriver à une fonction en un seul point ne va pas influencer la fonction lorsqu'elle est vue dans  $L^p$ . En tout cas lorsqu'on considère des mesures pour lesquelles les singletons sont de mesure nulle, et c'est bien le cas de la mesure de Lebesgue. Il est peut-être intuitivement moins clair que l'on peut non seulement modifier le comportement d'une fonction en un point, mais également modifier l'ensemble de base. En voici un exemple.

**Proposition 27.35.**

Nous avons les égalités suivantes d'espaces

$$L^p([0, 2\pi[) = L^p([0, 2\pi]) = L^p(S^1) \quad (27.85)$$

au sens où il existe des bijections isométriques de l'un à l'autre. Ici nous sous-entendons la mesure de Lebesgue partout<sup>15</sup>.

*Démonstration.* Voici une application où le crochet dénote la prise de classe :

$$\begin{aligned} \psi: L^p([0, 2\pi[) &\rightarrow L^p([0, 2\pi]) \\ [f] &\mapsto \text{la classe de } f_e(x) = \begin{cases} f(x) & \text{si } x \in ]0, 2\pi[ \\ 0 & \text{si } x = 0 \text{ ou } x = 2\pi. \end{cases} \end{aligned} \quad (27.86)$$

- (1) **Bien définie** Si  $[f] = [g]$  dans  $L^p([0, 2\pi[)$  alors  $f_e(x) = g_e(x)$  pour tout  $x \in ]0, 2\pi[$  sauf une partie de mesure nulle. L'union de cette partie avec  $\{0, 2\pi\}$  est encore de mesure nulle dans  $[0, 2\pi]$ . Les images par  $\psi$  sont donc égales dans  $L^p([0, 2\pi])$ .

15. Vu que la mesure de Lebesgue est définie pour  $\mathbb{R}^d$  munie de sa tribu des boréliens (complétée), vous êtes en droit de vous demander quelle est la tribu et la mesure que nous considérons sur le cercle  $S^1$ .

- (2) **Injective** Supposons que  $[f], [g] \in L^p(]0, 2\pi[)$  soient tels que  $\psi([f]) = \psi([g])$ . En particulier  $[f_e] = [g_e]$  et donc  $f_e = g_e$  presque partout sur  $[0, 2\pi]$ . Vu que  $\{0\} \cup \{2\pi\}$  est de mesure nulle, nous avons  $f_e = g_e$  presque partout sur  $]0, 2\pi[$ . Mais pour  $x \in ]0, 2\pi[$ , nous avons  $f_e(x) = f(x)$  et  $g_e(x) = g(x)$ . Bref nous avons  $f(x) = g(x)$  pour presque tout  $x \in ]0, 2\pi[$ .
- (3) **Surjective** Un élément de  $L^p([0, 2\pi])$  est l'image de sa restriction ... ou plutôt l'image de la classe de la restriction d'un quelconque de ses représentants.
- (4) **Isométrie** L'intégrale qui donne la norme sur  $L^p$  ne change pas selon que nous ajoutons ou non les bornes au domaine d'intégration.

De la même manière nous avons

$$L^p([0, 2\pi]) = L^p([0, 2\pi]). \quad (27.87)$$

En ce qui concerne l'identification avec  $L^p(S^1)$ , il faut passer par l'isométrie  $\varphi: [0, 2\pi[ \rightarrow S^1$  donnée par  $\varphi(t) = e^{it}$ , et être heureux que ce soit bien une isométrie parce qu'il faudra l'utiliser pour un changement de variables pour montrer que

$$\int_0^{2\pi} f(t) dt = \int_{S^1} (f \circ \varphi^{-1})(z) dz. \quad (27.88)$$

□

### 27.3.5 Inégalité de Young, Jensen, Hölder et de Minkowski

PROPOCCUBOOTMSI

**Proposition 27.36** (Inégalité de Young[720, 721]).

Soient  $a, b \geq 0$  ainsi que  $p, q > 0$  tels que

$$\frac{1}{p} + \frac{1}{q} = 1. \quad (27.89)$$

Alors

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \quad (27.90)$$

Il y a égalité si et seulement si  $a^p = b^q$ .

*Démonstration.* Nous fixons  $b \geq 0$ , et nous considérons la fonction

$$\begin{aligned} f: \mathbb{R}^+ &\rightarrow \mathbb{R} \\ x &\mapsto \frac{x^p}{p} + \frac{b^q}{q} - xb. \end{aligned} \quad (27.91)$$

On remarque vite fait que  $f'(x) = x^{p-1} - b$ . Nous avons  $f'(x) > 0$  si et seulement si  $x^{p-1} \geq b$ . Nous posons  $x_0 = b^{1/(p-1)}$ , et nous savons que  $f$  est strictement décroissante sur  $[0, x_0]$  et strictement croissante sur  $[x_0, \infty[$ . Cette fonction a donc un minimum global en  $x = x_0$ . Rapide calcul :

$$f(x_0) = \frac{b^{p/(p-1)}}{p} + \frac{b^q}{q} - b^{1+1/(p-1)}. \quad (27.92)$$

En utilisant  $q = p(p-1)$  nous trouvons vite que

$$f(x_0) = b^q \left( \frac{1}{p} + \frac{1}{q} - 1 \right) = 0. \quad (27.93)$$

Donc pour tout  $x \neq x_0$  nous avons  $f(x) > f(x_0) = 0$ , c'est-à-dire

$$\frac{a^p}{p} + \frac{b^q}{q} \geq ab \quad (27.94)$$

avec égalité si et seulement si  $a = x_0$ , c'est-à-dire quand  $a = b^{1/(p-1)}$  ou encore, en passant à la puissance  $p$  :

$$a^p = b^{p/(p-1)} = b^q. \quad (27.95)$$

□

PropXISooBxdaLk

**Proposition 27.37** (Inégalité de Jensen<sup>[523]</sup>).

Soit un espace mesuré de probabilité<sup>16</sup>  $(\Omega, \mathcal{A}, \mu)$  ainsi qu'une fonction convexe  $f: \mathbb{R} \rightarrow \mathbb{R}$  et une application  $\alpha: \Omega \rightarrow \mathbb{R}$  tels que  $\alpha$  et  $f \circ \alpha$  soient intégrables sur  $\Omega$ . Alors

$$f\left(\int_{\Omega} \alpha d\mu\right) \leq \int_{\Omega} (f \circ \alpha) d\mu. \quad (27.96)$$

*Démonstration.* Soient  $a \in \mathbb{R}$  et  $c_a$  le nombre donné par la proposition 17.90 : pour tout  $\omega \in \Omega$  nous avons

$$f(\alpha(\omega)) - f(a) \geq c_a(\alpha(\omega) - a). \quad (27.97)$$

EqOUMooLwknJP

Cela est en particulier vrai pour  $a = \int_{\Omega} \alpha d\mu$ . Nous intégrons l'inégalité (27.97) sur  $\Omega$  en nous souvenant que  $\int d\mu = 1$  :

$$\int_{\Omega} (f \circ \alpha) d\mu - \int_{\Omega} f(a) d\mu \geq c_a \left( \int_{\Omega} \alpha - \int_{\Omega} a \right) \quad (27.98a)$$

$$\int_{\Omega} (f \circ \alpha) d\mu - f(a) \geq c_a(a - a) \quad (27.98b)$$

$$f(a) \leq \int_{\Omega} (f \circ \alpha) d\mu. \quad (27.98c)$$

Cette dernière inégalité est celle que nous devons prouver.  $\square$

CORooRYUAooOHRsjL

**Corolaire 27.38.**

Soit un espace mesuré de probabilité  $(\Omega, \mathcal{A}, \mu)$  et une application  $\alpha \in L^1(\Omega, \mu)$  et  $\alpha \in L^p(\Omega)$  avec  $1 \leq p < +\infty$ . Alors

$$\left| \int_{\Omega} \alpha(s) d\mu(s) \right| \leq \|\alpha\|_p. \quad (27.99)$$

*Démonstration.* Il suffit d'utiliser l'inégalité de Jensen sur la fonction convexe  $f(x) = |x|^p$ . Nous avons alors

$$\left| \int_{\Omega} \alpha(s) d\mu(s) \right|^p \leq \int_{\Omega} |\alpha(s)|^p d\mu(s), \quad (27.100)$$

c'est-à-dire

$$\left| \int_{\Omega} \alpha(s) d\mu(s) \right| \leq \left[ \int_{\Omega} |\alpha(s)|^p d\mu(s) \right]^{1/p} = \|\alpha\|_p \quad (27.101)$$

où ma norme  $\|\cdot\|_p$  est prise au sens de la mesure  $\mu$ .  $\square$

ProptYqspT

**Proposition 27.39** (Inégalité de Hölder<sup>[722]</sup>).

Soit  $(\Omega, \mathcal{A}, \mu)$  un espace mesuré et  $1 \leq p, q \leq \infty$  satisfaisant  $\frac{1}{p} + \frac{1}{q} = 1$ . Si  $f \in L^p(\Omega)$ ,  $g \in L^q(\Omega)$ , alors nous avons les choses suivantes :

ITEMooNDKPooRKdmgS

(1) Le produit  $fg$  est dans  $L^1(\Omega)$  et nous avons

$$\|fg\|_1 \leq \|f\|_p \|g\|_q. \quad (27.102)$$

EqLPKooPBCQYN

(2) Si  $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$  alors

$$\|fg\|_r \leq \|f\|_p \|g\|_q \quad (27.103)$$

ITEMooQHLPOoRWWMOP

EqAVZooENyzmT

ITEMooBOYJooRkiAqJ

(3) Nous avons  $\|fg\|_1 = \|f\|_p \|g\|_q$  si et seulement si nous sommes dans un des trois cas suivants :

- $f = 0$  presque partout,
- $g = 0$  presque partout,
- Il existe  $\lambda > 0$  tel que  $|f|^p = \lambda |g|^q$  presque partout.

Ce qui est appelé « inégalité de Hölder » est généralement l'inéquation (27.102).

16. C'est-à-dire que  $\int_{\Omega} d\mu = 1$ .

*Démonstration.* En plusieurs points.

- (1) **Pour (1)** Nous allons le voir comme cas particulier de (2).  
 (2) **Pour (2)** D'abord nous supposons  $\|g\|_q = 1$  et nous posons

$$A = \{x \in \Omega \text{ tel que } |g(x)| > 0\}. \quad (27.104)$$

Hors de  $A$ , les intégrales que nous allons écrire sont nulles. Nous avons

$$\|fg\|_r^p = \left| \int_A |f|^r |g|^{r-q} |g|^q \right|^{p/r}, \quad (27.105)$$

et le coup tordu est de considérer cette intégrale comme étant une intégrale par rapport à la mesure  $\nu = |g|^q d\mu$  qui a la propriété d'être une mesure de probabilité par hypothèse sur  $g$ . Nous pouvons alors utiliser l'inégalité de Jensen<sup>17</sup> parce que  $p/r > 1$ , ce qui fait de  $x \mapsto |x|^{p/r}$  une fonction convexe. Nous avons alors

$$\|fg\|_r^p \leq \int_A (|f|^r |g|^{r-q})^{p/r} |g|^q d\mu \quad (27.106a)$$

$$= \int_A |f|^p |g|^{p(r-q)/r} |g|^q d\mu \quad (27.106b)$$

La puissance de  $|g|$  dans cette expression est :  $q + \frac{p(r-q)}{r} = 0$  parce que  $p(q-r) = rq$ . Nous avons alors montré que

$$\|fg\|_r^p \leq \int_A |f|^p d\mu \leq \|f\|_p^p. \quad (27.107)$$

La dernière inégalité est le fait que le domaine  $A$  n'est pas tout le domaine  $\Omega$ .

Si maintenant  $\|g\|_q \neq 1$  alors nous calculons

$$\|fg\|_r = \|g\|_q \|f \frac{g}{\|g\|_q}\|_r \leq \|g\|_q \|f\|_p \quad (27.108)$$

en appliquant la première partie à la fonction  $\frac{g}{\|g\|_q}$  qui est de norme 1.

- (3) **Pour (3)** Si  $f$  ou  $g$  est nulle presque partout, il y a immédiatement égalité. Nous supposons donc que  $f$  et  $g$  ne sont pas nulles presque partout et donc que  $\|f\|_p$  et  $\|g\|_q$  sont non nuls.

(3a) **Deux fonctions intermédiaires** Nous posons

$$\hat{f} = \frac{|f|}{\|f\|_p}, \quad \hat{g} = \frac{|g|}{\|g\|_q}. \quad (27.109)$$

(3b) **Égalité préliminaire** Nous avons

$$\frac{1}{p} \int \hat{f}^p = \frac{1}{p} \int \frac{|f|^p}{\|f\|_p^p} = \frac{1}{p\|f\|_p^p} \int |f|^p = \frac{1}{p}, \quad (27.110)$$

et de même avec  $g$  et  $q$  au lieu de  $f$  et  $p$ . Nous avons donc

$$\frac{1}{p} \int \hat{f}^p + \frac{1}{q} \int \hat{g}^q = 1. \quad (27.111)$$

(3c) **Les équivalences** Les choses suivantes sont équivalentes.

(3c.i)  $\int |fg| = \|f\|_p \|g\|_q$

(3c.ii)  $\int \hat{f} \hat{g} = 1$

(3c.iii)

$$\int \hat{f} \hat{g} = \frac{1}{p} \int \hat{f}^p + \frac{1}{q} \int \hat{g}^q \quad (27.112)$$

---

17. Proposition 27.37.

(3c.iv)

$$\int \left( \hat{g}\hat{f} - \frac{1}{p}\hat{f}^p - \frac{1}{q}\hat{g}^q \right) = 0. \quad (27.113)$$

(3c.v)

$$\hat{g}\hat{f} - \frac{1}{p}\hat{f}^p - \frac{1}{q}\hat{g}^q = 0 \quad (27.114)$$

presque partout.

En effet l'inégalité de Young <sup>18</sup> dit que l'intégrande est positive partout. Pour que l'intégrale soit nulle, il faut que l'intégrande soit nulle ; c'est le lemme 14.199.

(3c.vi)  $\hat{f}(x)^p = \hat{g}(x)^q$  pour presque tout  $x$ . C'est le cas d'égalité dans l'inégalité de Young.

(3c.vii)

$$\frac{|f|^p}{\|f\|_p^p} = \frac{|g|^q}{\|g\|_q^q}. \quad (27.115)$$

(3c.viii)

$$|f|^p = \lambda |g|^q \quad (27.116)$$

avec  $\lambda = \|f\|_p^p / \|g\|_q^q$ .

(3d) **Conclusion** En lisant les implications de haut en bas, nous avons la condition nécessaire au cas d'égalité. Pour traiter la condition suffisante, nous supposons qu'il existe  $\lambda > 0$  tel que  $|f|^p = \lambda |g|^q$ . Alors nous avons

$$\|f\|_p^p = \int |f|^p = \lambda \int |g|^q = \lambda \|g\|_q^q, \quad (27.117)$$

ce qui donne immédiatement  $\lambda = \|f\|_p^p / \|g\|_q^q$ . Nous pouvons donc remonter les équivalences.

□

RemNormuptNird

#### Remarque 27.40.

Dans le cas d'un espace de probabilité, la fonction constante  $g = 1$  appartient à  $L^p(\Omega)$ . En prenant  $p = q = 2$  nous obtenons

$$\|f\|_1 \leq \|f\|_2. \quad (27.118)$$

PROPooDVCCooKqbAwo

#### Proposition 27.41.

Soit un espace mesuré  $(\Omega, \mathcal{A}, \mu)$  vérifiant  $\mu(\Omega) = 1$ . Soient des nombres  $0 < r \leq s$  ainsi qu'une application  $f: \Omega \rightarrow \mathbb{R}$ .

Si  $\int_{\Omega} |f|^2 < \infty$ , alors  $\int_{\Omega} |f|^r < \infty$ .

*Démonstration.* Si  $s = r$ , la conclusions est immédiate. Nous supposons donc  $r \neq s$ . Nous considérons  $q$  tel que  $\frac{1}{s} + \frac{1}{q} = \frac{1}{r}$ , c'est-à-dire  $q = rs/(s-r)$ . Vu que  $s > r$ , nous avons  $q > 0$ , pas de problème.

Par hypothèse, nous avons  $f \in L^s(\Omega)$ . Et comme  $\Omega$  est de mesure finie, la fonction constante  $1_{\Omega}$  est dans  $L^q(\Omega)$ . Nous pouvons donc écrire l'inégalité de Hölder 27.103 :

$$\|f 1_{\Omega}\|_r \leq \|f\|_s \|1_{\Omega}\|_q. \quad (27.119)$$

Cela donne  $\|f\|_r \leq \|f\|_s$ , et en déballant,

$$\left( \int_{\Omega} |f|^r \right)^{1/r} \leq \left( \int_{\Omega} |f|^s \right)^{1/s}. \quad (27.120)$$

Vu que  $r > 0$ , l'application  $t \mapsto t^r$  est croissante, et nous pouvons écrire

$$\int_{\Omega} |f|^r \leq \left( \int_{\Omega} |f|^s \right)^{r/s} < \infty. \quad (27.121)$$

Notez l'utilisation de la proposition 12.432.

□

18. Proposition 27.36.

LemTLHwYzD

**Lemme 27.42.**À propos de  $L^1_{loc}(I)$ .

- (1) Lorsque  $I$  est borné nous avons  $L^2(I) \subset L^1(I)$ .  
 (2) Si  $I$  n'est pas borné alors  $L^2(I) \subset L^1_{loc}(I)$ .

*Démonstration.* En effet si  $I$  est borné, alors la fonction constante 1 est dans  $L^2(I)$  et l'inégalité de Hölder 27.39 nous dit que le produit  $1u$  est dans  $L^1(I)$ .

Si  $I$  n'est pas borné, nous refaisons le même raisonnement sur un compact  $K$  de  $I$ .  $\square$

CORooIEAooNmbkTo

**Corolaire 27.43** ([723]).

Soit l'espace  $L^2(I)$  avec  $I = ]0, 1[$  avec la mesure de Lebesgue. Si  $u_n \in L^2$  converge vers  $u$  dans  $L^2$  alors nous pouvons permuter l'intégrale et la limite :

$$\lim_{n \rightarrow \infty} \int_I u_n = \int_I u. \quad (27.122)$$

*Démonstration.* Nous considérons la forme linéaire

$$\begin{aligned} T: L^2(I) &\rightarrow \mathbb{R} \\ u &\mapsto \int_I u. \end{aligned} \quad (27.123)$$

Elle est bien définie par l'inégalité de Hölder  $\|fg\|_1 \leq \|f\|_2 \|g\|_2$  appliqué à  $g(x) = 1$  qui vérifie  $\|g\|_2 = 1$ . Nous avons aussi

$$T(u) \leq \int_I |u| \leq \|u\|_1 \leq \|u\|_2 \quad (27.124)$$

où la dernière inégalité est celle de Hölder 27.39. Bref,  $T$  est continue. Cela signifie que si  $u_n \xrightarrow{L^2(I)} u$  alors  $T(u_n) = T(u)$ . Cela est l'égalité demandée.  $\square$

**27.3.6 Inégalité de Minkowski****27.3.6.1 Forme usuelle****⚠ Avertissement/question au lecteur !! 27.44**

Dans la proposition suivante, la partie « égalité » est très personnelle. Je n'en n'ai pas trouvé de preuve complète. Donc soyez doublement vigilant et écrivez-moi si vous avez quelque chose à dire.

PropInegMinkKUpRHg

**Proposition 27.45** (Inégalité de Minkowski[724, 725, 726, 727, 728, 1]).

Si  $1 \leq p < \infty$  et si  $f, g \in L^p(\Omega, \mathcal{A}, \mu)$  alors

ItemDHukLJi

- (1)  $\|f + g\|_p \leq \|f\|_p + \|g\|_p$   
 (2) Si  $p = 1$ , il y a égalité si et seulement si  $f(x)\overline{g(x)} \geq 0$  pour presque tout  $x$ .  
 (3) Si  $p > 1$ , il y a égalité si et seulement si il existe des réels positifs  $\alpha, \beta$  pas tous deux nuls tels que

ITEMooGRXBooMLRMww

ITEMooQCSHooNUDwtM

$$\alpha f(x) = \beta g(x) \quad (27.125)$$

pour presque tout  $x$ .

*Démonstration.* En plusieurs points.

- (1) **Pour (1)** En utilisant l'inégalité  $|z_1 + z_2| \leq |z_1| + |z_2|$  (proposition 9.108(1)) pour chaque  $x$  dans l'intégrale, de façon à pouvoir majorer

$$|f(x) + g(x)|^p = |f(x) + g(x)| |f(x) + g(x)|^{p-1} \quad (27.126a)$$

$$\leq (|f(x)| + |g(x)|) |f(x) + g(x)|^{p-1}. \quad (27.126b)$$

Nous mettons ça dans une intégrale et nous calculons un peu : SUBEQSooGWMTooDBXSgL

$$\|f + g\|_p^p = \int |f + g|^p d\mu \quad (27.127a)$$

$$\leq \int (|f| + |g|) |f + g|^{p-1} d\mu \quad (27.127b) \quad \text{SUBEQooYCTWooQpHNqa}$$

$$= \int |f| |f + g|^{p-1} + \int |g| |f + g|^{p-1} \quad (27.127c) \quad \text{SUBEQooNQKBooQFtTH}$$

Lorsque  $p = 1$ , nous nous arrêtons ici parce que (27.127c) s'écrit

$$\|f + g\|_1 \leq \int |f| + \int |g| = \|f\|_1 + \|g\|_1. \quad (27.128)$$

Lorsque  $p > 1$ , nous devons continuer et utiliser Hölder. Attardons nous sur le premier terme. Nous posons  $q = p/(p-1)$ ,  $a = f$  et  $b = |f + g|^{p-1}$ , et nous utilisons l'inégalité de Hölder <sup>19</sup> : SUBEQSooFINUooQfdMS

$$\int |f| |f + g|^{p-1} = \|ab\|_1 \quad (27.129a)$$

$$\leq \|a\|_p \|b\|_q \quad (27.129b)$$

$$= \|f\|_p \left[ \int (|f + g|^{p-1})^q \right]^{1/q} \quad (27.129c)$$

$$= \|f\|_p \left( \int |f + g|^p \right)^{1-\frac{1}{p}} \quad (27.129d)$$

$$= \|f\|_p \left( \int |f + g|^p \right)^{(p-1)/p} \quad (27.129e)$$

$$= \|f\|_p \|f + g\|_p^{p-1} \quad (27.129f)$$

Nous avons utilisé la règle de produit d'exposants et de somme d'exposants <sup>20</sup>

Nous utilisons cette inégalité dans les deux termes de (27.127) : SUBEQSooKHKOooOCfndf

$$\|f + g\|_p^p \leq \int |f| |f + g|^{p-1} + \int |g| |f + g|^{p-1} \quad (27.130a)$$

$$\leq \|f\|_p \|f + g\|_p^{p-1} + \|g\|_p \|f + g\|_p^{p-1} \quad (27.130b)$$

Nous obtenons le résultat en divisant le tout par  $\|f + g\|_p^{p-1}$ .

- (2) **Pour (2)** Toutes les inégalités de (27.127) sont des égalités. En particulier nous avons celle-ci :

$$\int |f + g| d\mu = \int (|f| + |g|) d\mu \quad (27.131)$$

Cela donne

$$\int (|f + g| - |f| - |g|) d\mu = 0. \quad (27.132)$$

Vue l'inégalité de la proposition 9.108(1), la fonction intégrée est toujours négative. Pour que l'intégrale soit nulle, il faut que la fonction intégrée soit presque partout nulle :

$$|f(x) + g(x)| = |f(x)| + |g(x)| \quad (27.133)$$

pour presque tout  $x$ . La partie « égalité » de la proposition 9.108(1) donne alors le résultat.

\_\_\_\_\_ Cas d'égalité, préliminaire \_\_\_\_\_

19. Proposition 27.39.

20. Propositions 12.432 et 12.422(3).

Nous commençons par prouver que si  $f + g = 0$  alors  $f = g = 0$ . Pour cela nous faisons le calcul

$$\begin{aligned}
 \|f\|_p + \|g\|_p &= \|f + g\| && \text{SUEQooPUQGooAIsFFa} \\
 &\leq \| |f| + |g| \| && \text{(27.134a)} \\
 &\leq \| |f| \|_p + \| |g| \| && \text{SUEQooFXHBooYXMBdf} \\
 &= \|f\|_p + \|g\|_p. && \text{(27.134b)} \\
 & && \text{SUEQooQZFXooPYgQrI} \\
 & && \text{(27.134c)} \\
 & && \text{(27.134d)}
 \end{aligned}$$

Justifications :

- Pour (27.134a). L'hypothèse d'égalité dans l'inégalité de Minkowski.
- Pour (27.134b). Intégrale de l'inégalité de la proposition 9.108(1).
- Pour (27.134c). Inégalité de Minkowski.

Étant donné que le premier et le dernier membre de (27.134) sont égaux, toutes les inégalités sont des égalités. En particulier,

$$\|f + g\|_p = \| |f| + |g| \|_p. \quad (27.135)$$

En termes d'intégrales, cette égalité donne

$$\int |f + g|^p - (|f| + |g|)^p = 0. \quad \text{EQooRYQBooGdgtLN} \quad (27.136)$$

L'inégalité déjà mentionnée  $|z_1 + z_2|^p \leq (|z_1| + |z_2|)^p$  dit que la fonction à intégrer dans (27.136) est partout négative. Pour que l'intégrale soit nulle, il faut donc que la fonction soit presque partout nulle. Autrement dit, pour presque tout  $x$  nous avons

$$|f(x) + g(x)| = |f(x)| + |g(x)|. \quad \text{EQooULHAooWXVIEH} \quad (27.137)$$

Cette égalité nous sera utile dans la suite. En particulier, reprenez que si  $f + g = 0$  presque partout, alors  $f = g = 0$  presque partout. Dans ce cas, l'égalité  $\alpha f + \beta g = 0$  fonctionne pour tout  $\alpha$  et  $\beta$ .

\_\_\_\_\_ Cas d'égalité (2),  $p = 1$ . \_\_\_\_\_

(1)  $\Rightarrow$

L'inégalité (27.127b) est une égalité :

$$\int |f + g|^{p-1} (|f + g| - |f| - |g|) = 0. \quad (27.138)$$

Par la proposition 9.108(1), la fonction intégrée est presque partout nulle ou positive. Pour que l'intégrale soit nulle, il faut que la fonction soit presque partout nulle (lemme 14.199). Presque chaque  $x$  vérifie sont une des deux conditions suivantes :

$$(1a) \quad |f(x) + g(x)| = 0$$

$$(1b) \quad |f(x) + g(x)| - |f(x)| - |g(x)| = 0.$$

Pour les  $x$  vérifiant la première condition, nous avons déjà vu que  $f = g = 0$ , juste en dessous de (27.137). En particulier, nous avons  $f(x)\overline{g(x)} \geq 0$  pour tout  $x$  dans ce cas.

En ce qui concerne les  $x$  vérifiant la seconde condition, la proposition 9.108(1) montre qu'ils vérifient  $f(x)\overline{g(x)} \geq 0$ .

Pour les  $x$  dans le second cas, la proposition 9.108(1) donne  $f(x)\overline{g(x)} \geq 0$ .

Bref, dans tous les cas nous avons  $f(x)\overline{g(x)} \geq 0$  pour presque tout  $x$ .

(2)  $\Leftarrow$

Nous supposons que  $f(x)\overline{g(x)} \geq 0$  pour presque tout  $x$ . La proposition 9.108(1) (utilisée dans l'autre sens) dit que  $|f(x) + g(x)| - |f(x)| - |g(x)| = 0$  pour presque tout  $x$ . L'inégalité (27.127b) est alors une égalité.



————— Cas d'égalité (3),  $p > 1$ . —————

Il y a deux sens.

- (1)  $\Rightarrow$ , les inégalités Nous supposons que  $f, g \in L^p(\Omega, \mathcal{A}, \mu)$  vérifient  $\|f + g\|_p = \|f\|_p + \|g\|_p$ , et nous allons voir ce que cela implique. Regroupons les différentes inégalités qui mènent à Minkowski.

$$\begin{aligned}
 \|f + g\|_p^p &= \int |f + g|^p d\mu & (27.139a) \\
 &\leq \int (|f| + |g|)|f + g|^{p-1} d\mu & \text{eq. (27.127)} \quad \text{SUBEQooXPMYooJBKwZH} \quad (27.139b) \\
 &\leq \|f\|_p \|f + g\|_p^{p-1} + \int |g| |f + g|^{p-1} d\mu & \text{eq. (27.130)} \quad \text{SUBEQooUPPAooIleVbN} \quad (27.139c) \\
 &\leq \|f\|_p \|f + g\|_p^{p-1} + \|g\|_p \|f + g\|_p^{p-1} & \text{eq. (27.130)} \quad (27.139d) \\
 &= (\|f\|_p + \|g\|_p) \|f + g\|_p^{p-1} & (27.139e) \\
 &= \|f + g\|_p^p & \text{hyp. égalité.} \quad (27.139f)
 \end{aligned}$$

La dernière ligne est l'hypothèse d'égalité  $\|f\|_p + \|g\|_p = \|f + g\|_p$ . Comme le premier et le dernier membre de ces inégalités sont égaux, toutes les inégalités sont des égalités.

- (2)  $\Rightarrow$ , première inégalité

L'inégalité (27.139b) est une égalité :

$$\int |f + g|^{p-1} (|f + g| - |f| - |g|) = 0. \quad (27.140)$$

Par la proposition 9.108(1), la fonction intégrée est presque partout nulle ou positive. Pour que l'intégrale soit nulle, il faut que la fonction soit presque partout nulle (lemme 14.199). Presque chaque  $x$  vérifie sont une des deux conditions suivantes :

$$(2a) \quad |f(x) + g(x)| = 0$$

$$(2b) \quad |f(x) + g(x)| - |f(x)| - |g(x)| = 0.$$

Pour les  $x$  vérifiant la première condition, nous avons déjà vu que  $\alpha f + \beta g = 0$  pour n'importe quels  $\alpha$  et  $\beta$ , juste en dessous de (27.137).

En ce qui concerne les  $x$  vérifiant la seconde condition, la proposition 9.108(1) montre qu'ils vérifient  $f(x)\overline{g(x)} \geq 0$ .

Nous avons donc déjà montré que si  $f$  et  $g$  vérifient une égalité, alors pour presque tout  $x$ , nous avons

$$f(x)\overline{g(x)} \geq 0. \quad (27.141)$$

- (3)  $\Rightarrow$ , seconde inégalité (Hölder)

La seconde inégalité à être une égalité est (27.139c) :

$$\int |f| |f + g|^{p-1} = \|f\|_p \|f + g\|_p^{p-1}. \quad (27.142)$$

Il s'agit du cas d'égalité de Hölder avec  $a = f$ ,  $b = |f + g|^{p-1}$  traité dans la proposition 27.39(3). Nous sommes donc dans un des trois cas suivants :

- $f = 0$  presque partout,
- $f + g = 0$  presque partout,
- il existe  $\lambda > 0$  tel que  $|a|^p = \lambda |b|^q$ .

Dans le premier cas, nous avons  $\alpha f = \beta g$  avec  $\alpha = 1$  et  $\beta = 0$ . Le second cas donne  $f = g = 0$  et donc  $\alpha f + \beta g = 0$  pour n'importe quels  $\alpha$  et  $\beta$  (voir autour de (27.137)).

Nous supposons donc être dans le troisième cas. Étant donné que  $q = p/(p-1)$ , nous avons

$$|b|^q = (|f + g|^{p-1})^q = |f + g|^p, \quad (27.143)$$

et donc la contrainte est l'existence de  $\lambda > 0$  tel que  $|f|^p = \lambda |f + g|^p$ .

(4)  $\Rightarrow$ , troisième inégalité (Hölder)

La seconde application de Hölder est avec  $a = g$  et  $b = |f + g|^{p-1}$ . Nous sommes dans un cas d'égalité si nous sommes dans un des trois cas suivants :

- $g = 0$
- $f + g = 0$
- il existe  $\sigma > 0$  tel que  $|a|^p = \sigma|b|^q$ .

Encore une fois nous supposons être dans le troisième cas. Il existe donc  $\sigma > 0$  tel que  $|g|^p = \sigma|f + g|^p$ .

(5) Petit résumé

Nous avons prouvé l'existence de  $\lambda, \sigma > 0$  tels que

$$\begin{cases} f(x)\overline{g(x)} \geq 0 & \text{pour presque tout } x & (27.144a) \\ |f|^p = \lambda|f + g|^p & & (27.144b) \\ |g|^p = \sigma|f + g|^p & & (27.144c) \end{cases}$$

(6) Des calculs et des cas En passant à la racine  $p^e$  :

$$\begin{cases} f(x)\overline{g(x)} \geq 0 & \text{pour presque tout } x & (27.145a) \\ |f| = s|f + g| & & (27.145b) \\ |g| = t|f + g|. & & (27.145c) \end{cases}$$

Les deux dernières égalités du système permettent d'écrire

$$t|f| = st|f + g| = s|g|. \quad (27.146)$$

Nous considérons donc le système, valide pour presque tout  $x$  : SUBEQSooLKGDooQouvhW

$$\begin{cases} f(x)\overline{g(x)} \geq 0 & (27.147a) \\ t|f(x)| = s|g(x)| & (27.147b) \end{cases}$$

Avec la contrainte  $(s, t) \neq (0, 0)$ .

La proposition 9.107 appliqué à chaque couple  $f(x), g(x)$  donne l'existence d'une fonction réelle  $a$  telle que pour chaque  $x$  nous ayons un des deux cas suivants :

- $g(x) = 0$
- $f(x) = a(x)g(x)$ .

De plus  $a(x) \in \mathbb{R}^+$  parce que  $f(x)\overline{g(x)}$  est non seulement réel, mais également positif.

(6a)  $s = 0$  Alors nous avons  $t|f(x)| = 0$  et donc  $f(x) = 0$  pour presque tout  $x$  parce que  $t \neq 0$ . Nous avons alors  $\alpha f(x) + \beta g(x) = 0$  avec  $\alpha = 1$  et  $\beta = 0$ .

(6b)  $s \neq 0$  Les points de  $\Omega$  se séparent en deux parties :  $g(x) = 0$  et les autres.

(6b.i)  $g(x) \neq 0$  Dans ce cas nous avons le système

$$\begin{cases} f(x) = a(x)g(x) & (27.148a) \\ t|f(x)| = s|g(x)|. & (27.148b) \end{cases}$$

Vu que  $s \neq 0$  et  $g(x) \neq 0$ , la dernière équation donne  $t|f(x)| \neq 0$  et donc  $t \neq 0$  et  $f(x) \neq 0$ .

La première équation donne  $a(x) = f(x)/g(x)$  que nous pouvons mettre dans la secondes pour obtenir :

$$|a(x)| = \frac{s}{t}. \quad (27.149)$$

Vu que  $a$  est à valeurs dans  $\mathbb{R}^+$ , cela donne

$$a(x) = \frac{s}{t} \quad (27.150)$$

pour tout  $x$  tel que  $g(x) \neq 0$ .

(6b.ii)  $g(x) = 0$  Dans ce cas, le système (27.147) devient

$$t|f(x)| = 0. \quad (27.151)$$

Vu que  $s \neq 0$ , nous avons  $f(x) = 0$ . Donc n'importe quel choix de  $\alpha$  et  $\beta$  fait l'affaire pour ces points. Il n'y a pas de contraintes.

(6c)  $\Leftarrow$  Nous supposons  $\alpha f(x) = \beta g(x)$  pour presque tout  $x$ . Pour fixer les idées, nous supposons que  $\beta \neq 0$  (sinon, refaire le raisonnement en inversant les rôles de  $f$  et  $g$ ). En posant  $\lambda = \alpha/\beta$  nous avons

$$g(x) = \lambda f(x) \quad (27.152)$$

pour presque tout  $x$ . En passant par l'intégrale,

$$\|f + g\|_p = \|(1 + \lambda)f\|_p = (1 + \lambda)\|f\|_p = \|f\|_p + \lambda\|f\|_p = \|f\|_p + \|\lambda f\|_p = \|f\|_p + \|g\|_p. \quad (27.153)$$

Le  $\lambda$  et  $(1 + \lambda)$  rentrent et sortent des normes parce qu'ils sont positifs.

□

### 27.3.6.2 Inégalité intégrale de Minkowski

Dans le cas où nous n'avons pas une somme de deux fonctions mais d'une infinité paramétrée par  $y \in Y$ , nous pouvons convertir la somme de l'inégalité de Minkowski en une intégrale :

$$\left[ \int_X \left( \int_Y |f(x, y)| d\nu(y) \right)^p d\mu(x) \right]^{1/p} \leq \int_Y \left( \int_X |f(x, y)|^p d\mu(x) \right)^{1/p} d\nu(y). \quad (27.154)$$

Ce sera la proposition 27.50.

LEMooMZWYooIsJJms

**Lemme 27.46** ([1]).

Soit un espace mesuré  $X$ . Soient  $p, q$  tels que  $\frac{1}{p} + \frac{1}{q} = 1$ . Soient des applications  $f, g: X \rightarrow \mathbb{C}$  telles que  $f \in L^p(X)$  et  $g^{p-1} \in L^q(X)$ . Alors

$$\int_X f(x) g^{p-1}(x) dx \leq \|f\|_p \|g\|_p^{p-1}. \quad (27.155)$$

*Démonstration.* Nous utilisons Hölder 27.39(1) :

$$\int_X f(x) g^{p-1}(x) dx = \|f g^{p-1}\|_1 \leq \|f\|_p \|g^{p-1}\|_q. \quad (27.156)$$

Le seul truc à savoir est que  $q(p-1) = p$  et  $p/q = p-1$  ; à part ça, nous calculons

$$\|g^{p-1}\|_q = \left( \int_X |g^{p-1}(x)|^q dx \right)^{1/q} \quad (27.157a)$$

$$= \left( \int_X |g(x)|^{q(p-1)} dx \right)^{1/q} \quad (27.157b)$$

$$= \left( \int_X |g(x)|^p dx \right)^{1/q} \quad (27.157c)$$

$$= \left( \|g\|_p^p \right)^{1/q} = \|g\|_p^{p/q} = \|g\|_p^{p-1}. \quad (27.157d)$$

□

PROPooYOWUooXoNLbL

**Proposition 27.47** ([729, 1]).

Soit  $1 \leq p \leq \infty$  ainsi que  $q \in \mathbb{R}$  tel que  $\frac{1}{p} + \frac{1}{q} = 1$ . Soit un espace mesuré  $\sigma$ -fini  $(\Omega, \mathcal{A}, \mu)$ . Soit  $f: \Omega \rightarrow \mathbb{C}$ .

ITEMooONTQooMHHOxk

(1) Si  $p < \infty$  et si  $\|f\|_p < \infty$ , alors

$$\|f\|_p = \max_{\|\varphi\|_q \leq 1} \left| \int_{\Omega} f \varphi d\mu \right|. \quad (27.158)$$

ITEMooBZXAOoAfgWvF

(2) Si  $p \leq \infty$ , alors

$$\|f\|_p = \sup_{\|\varphi\|_q \leq 1} \left| \int_{\Omega} f \varphi d\mu \right|. \quad (27.159)$$

*Démonstration.* Notez l'inégalité de Hölder 27.39(1) :

$$\left| \int_{\Omega} f \varphi \right| \leq \int_{\Omega} |f \varphi| = \|f \varphi\|_1 \leq \|f\|_1 \|\varphi\|_q. \quad (27.160)$$

Donc l'inégalité dans un sens est toujours vérifiée. Nous ne nous concentrons que sur l'inégalité dans l'autre sens.

En plein de parties.

(1) **Pour (1)** Notons tout de suite qu'on suppose  $\|f\|_p \neq 0$ . Sinon  $f$  est presque partout nulle et n'importe que  $\varphi$  fait l'affaire. Nous subdivisons encore en deux parties.

(1a) Si  $\|f\|_p = 1$  Nous posons

$$\begin{aligned} \varphi: \Omega &\rightarrow \mathbb{C} \\ \omega &\mapsto \begin{cases} \frac{|f(\omega)|^p}{f(\omega)} & \text{si } f(\omega) \neq 0 \\ 0 & \text{sinon.} \end{cases} \end{aligned} \quad (27.161)$$

Nous avons bien  $\|\varphi\|_q = 1$  parce que, en remarquant que  $q(p-1) = 1$ ,

$$\|\varphi\|_q^q = \int_{\Omega} |\varphi(\omega)|^q d\omega = \int_{\Omega_1} (|f(\omega)|^{p-1})^q = \int_{\Omega} |f(\omega)|^p d\omega = 1. \quad (27.162)$$

(1b) Si  $\|f\|_p < \infty$  Nous posons  $g = f/\|f\|_p$ . Vu que  $\|g\|_p = 1$ , nous reprenons le  $\varphi$  du cas précédent qui va pour  $g$ . Et nous vérifions que ce  $\varphi$  fonctionne aussi pour  $f$  :

$$\int_{\Omega} f \varphi = \|f\|_p \int_{\Omega} g \varphi = \|f\|_p. \quad (27.163)$$

(2) **Pour (2), avec  $p < \infty$  et  $\|f\|_p = \infty$**  Nous supposons que  $p < \infty$  et que  $\|f\|_p = \infty$ . Il y a deux cas. Si  $f = \infty$  sur une partie de mesure non nulle, il suffit de prendre  $\varphi$  non nulle sur cette partie, et on aura facilement  $\int_{\Omega} f \varphi = \infty$ .

Nous supposons que  $|f| = \infty$  seulement sur une partie de mesure nulle, et nous considérons des parties  $\Omega_k$  comme dans la proposition 27.21. Ensuite nous posons

$$\begin{aligned} \alpha_k: \Omega &\rightarrow \mathbb{C} \\ \omega &\mapsto \begin{cases} 0 & \text{si } \omega \notin \Omega_k \\ 0 & \text{si } f(\omega) = 0 \\ \frac{|f(\omega)|^p}{f(\omega)} & \text{sinon} \end{cases} \end{aligned} \quad (27.164)$$

et  $\varphi_k = \alpha_k / \|\alpha_k\|_q$ . Nous avons manifestement  $\|\varphi_k\|_q = 1$ .

Commençons par calculer ce que vaut réellement  $\|\alpha_k\|_q$ . En notant  $\Omega'_k = \{\omega \in \Omega_k \text{ tel que } f(\omega) \neq 0\}$  nous avons un petit calcul :

$$\|\alpha_k\|_q = \left( \int_{\Omega} |\alpha_k(\omega)|^q \right)^{1/q} \quad (27.165a)$$

$$= \left( \int_{\Omega'_k} \frac{|f(\omega)|^{pq}}{|f(\omega)|^q} d\omega \right)^{1/q} \quad (27.165b)$$

$$= \left( \int_{\Omega'_k} |f(\omega)|^p \right)^{1/q} \quad (27.165c)$$

$$= \|f\|_{k,p}^{p/q} \quad (27.165d)$$

où nous avons noté  $\|f\|_{k,p}$  la norme  $L^p$  de  $f$  sur  $\Omega_k$ .

Nous pouvons maintenant prouver notre énoncé :

$$\int_{\Omega} f \varphi_k = \frac{1}{\|\alpha_k\|_q} \int_{\Omega'_k} f(\omega) \frac{|f(\omega)|^k}{f(\omega)} d\omega \quad (27.166a)$$

$$= \frac{1}{\|\alpha_k\|_q} \int_{\Omega_k} |f(\omega)|^p d\omega \quad (27.166b)$$

$$= \frac{1}{\|\alpha_k\|_q} \|f\|_{k,p}^p \quad (27.166c)$$

$$= \|f\|_{k,p}^{p-p/q} \quad (27.166d)$$

$$= \|f\|_{k,p}. \quad (27.166e)$$

La proposition 27.22 conclu que

$$\int_{\Omega} f \varphi_k = \|f\|_{k,p} \rightarrow \|f\|_p = \infty. \quad (27.167)$$

(3) Pour (2), avec  $p = \infty$  Nous passons au cas  $p = \infty$ <sup>21</sup>. Soit  $\epsilon \in ]0, 1[$ , et posons

$$A_{\epsilon} = \{\omega \in \Omega \text{ tel que } |f(\omega)| > 1 - \epsilon\}. \quad (27.168)$$

Notez que  $A_{\epsilon} \in \mathcal{A}$ .

(3a) **Si**  $\|f\|_{\infty} = 1$  Nous avons  $\mu(A_{\epsilon}) > 0$ , sinon nous aurions  $|f(x)| < 1 - \epsilon$  presque partout et donc  $\|f\|_{\infty} < 1$ . Vu que  $(\Omega, \mu)$  est  $\sigma$ -fini, il existe une partie mesurable  $B_{\epsilon} \subset A_{\epsilon}$  vérifiant  $0 < \mu(B_{\epsilon}) < \infty$ . Nous posons alors

$$\begin{aligned} \varphi_{\epsilon} : \Omega &\rightarrow \mathbb{C} \\ x &\mapsto \begin{cases} \frac{|f(x)|}{\mu(B_{\epsilon})f(x)} & \text{si } x \in B_{\epsilon} \\ 0 & \text{sinon.} \end{cases} \end{aligned} \quad (27.169)$$

Notons que  $B_{\epsilon} \subset A_{\epsilon}$  et donc  $f(x) \neq 0$  sur  $B_{\epsilon}$ . La fonction est donc bien définie (pas de problèmes avec le dénominateur).

Sur  $B_{\epsilon}$ , nous avons  $|\varphi_{\epsilon}(x)| = 1/\mu(B_{\epsilon}) < \infty$ . Donc  $\varphi_{\epsilon} \in L^1(\Omega)$  parce que  $B_{\epsilon}$  est de mesure finie. De plus

$$\int_{\Omega} |\varphi_{\epsilon}(x)| dx = \int_{B_{\epsilon}} \frac{1}{\mu(B_{\epsilon})} dx = 1. \quad (27.170)$$

Nous avons prouvé que pour tout  $\epsilon \in ]0, 1[$ , nous avons

$$\sup_{\|\varphi\|=1} \int_{\Omega} f \varphi \geq 1 - \epsilon, \quad (27.171)$$

21. Et c'est le moment de ne pas confondre  $\|f\|_{\infty}$  qui signifie en réalité  $\|f\|_{L^{\infty}}$  avec la norme suprénum qui n'a rien à voir avec que nous faisons ici.

et donc que

$$\sup_{\|\varphi\|=1} \int_{\Omega} f\varphi = 1 = \|f\|_{\infty}. \quad (27.172)$$

- (4) **Si**  $\|f\|_{\infty} \neq 1$  Nous posons  $g = f/\|f\|_{\infty}$ , de telle sorte que  $\|g\|_{\infty} = 1$ . Nous prenons les  $\varphi_{\epsilon}$  correspondant à  $g$  et nous avons

$$\int_{\Omega} f\varphi_{\epsilon} = \|f\|_{\infty} \int_{\Omega} g\varphi_{\epsilon} \geq \|f\|_{\infty}(1 - \epsilon), \quad (27.173)$$

et donc ça marche encore.

□

### 27.48.

Pour la suite, nous considérons deux espaces mesurés  $\sigma$ -finis  $X$  et  $Y$  ainsi qu'une application  $h: X \times Y \rightarrow \mathbb{C}$ . Nous notons alors

$$\begin{aligned} h_{1,y}: X &\rightarrow \mathbb{C} \\ x &\mapsto h(x, y), \end{aligned} \quad (27.174)$$

et

$$\begin{aligned} h_{2,x}: Y &\rightarrow \mathbb{C} \\ y &\mapsto h(x, y), \end{aligned} \quad (27.175)$$

et

$$\begin{aligned} N_1: Y &\rightarrow \mathbb{R} \\ y &\mapsto \|h_{1,y}\|_p, \end{aligned} \quad (27.176)$$

et

$$\begin{aligned} N_2: X &\rightarrow \mathbb{R} \\ x &\mapsto \|h_{2,x}\|_1 \end{aligned} \quad (27.177)$$

pour tous les  $p \in \mathbb{R}$  pour lesquels  $N_2$  a un sens.

PROPOoUMSFooJWFJzt

### Proposition 27.49 ([730]).

Soient deux espaces mesurés  $\sigma$ -finis  $X$  et  $Y$ . Soit  $1 \leq p < \infty$ . Soit une application mesurable  $h: X \times Y \rightarrow \mathbb{C}$ . Alors l'application

$$\begin{aligned} N_1: Y &\rightarrow \mathbb{R} \\ y &\mapsto \|h_{1,y}\|_p \end{aligned} \quad (27.178)$$

est mesurable.

*Démonstration.* Explicitons :

$$N(y) = \int_X |h(x, y)|^p dy. \quad (27.179)$$

Le théorème de Fubini-Tonelli 14.302(1) appliqué à la fonction

$$\begin{aligned} f: X \times Y &\rightarrow \mathbb{R} \\ (x, y) &\mapsto |h(x, y)|^p \end{aligned} \quad (27.180)$$

dit que  $N$  est mesurable.

□

PROPOoINXBBooTrTxwg

### Proposition 27.50 (Inégalité intégrale de Minkowski[1, 731, 732, 132]).

Soit  $1 \leq p \leq \infty$ . Soient deux espaces mesurables  $\sigma$ -finis  $X$  et  $Y$ . Soit  $h: X \times Y \rightarrow \mathbb{C}$ . Nous supposons :

- (1)  $h$  est mesurable.
- (2) Pour presque tout  $y \in Y$ , nous avons  $h_{1,y} \in L^p(X)$ .
- (3)  $N_1 \in L^1(Y)$ .

Alors nous avons :

(1)  $h_{2,x} \in L^1(Y)$  pour presque tout  $x \in X$ .

(2)  $N_2 \in L^p(X)$ .

(3) La formule :

$$\left( \int_X \left| \int_Y h(x, y) dy \right|^p dx \right)^{1/p} \leq \int_Y \left( \int_X |h(x, y)|^p dx \right)^{1/p} dy \quad (27.181)$$

ou, en posant

$$\begin{aligned} F: X &\rightarrow \mathbb{C} \\ x &\mapsto \int_Y h(x, y) dy, \end{aligned} \quad (27.182)$$

l'inégalité

$$\|F\|_p \leq \int_Y \|h_{1,y}\|_p dy. \quad (27.183)$$

Note : l'équation (27.183) est plus adaptée que (27.181) pour l'énoncé dans le cas  $p = \infty$ .

Démonstration. Nous subdivisons selon les valeurs de  $p$ . Et d'ailleurs, les points (1), (2) et (3) ne seront pas démontrés dans l'ordre.

(1) **Pour**  $p = 1$  Point par point.

(1a) **Pour** (3) En utilisant Fubini-Tonelli 14.302(2),

$$\int_X \left| \int_Y h(x, y) dy \right| dx \leq \int_X \left( \int_Y |h(x, y)| dy \right) dx = \int_Y \left( \int_X |h(x, y)| dx \right) dy. \quad (27.184)$$

(1b) **Pour** (1) En posant

$$\begin{aligned} F: X &\rightarrow [0, \infty] \\ x &\mapsto \|h_{2,x}\|_1, \end{aligned} \quad (27.185)$$

nous avons

$$F(x) = \|h_{2,x}\|_1 = \int_Y |h(x, y)| dy. \quad (27.186)$$

En permutant les intégrales par Fubini-Tonelli,

$$\int_X F(x) = \int_X \left( \int_Y |h(x, y)| dy \right) dx \quad (27.187a)$$

$$= \int_Y \left( \int_X |h(x, y)| dx \right) dy \quad (27.187b)$$

$$= \int_Y N_1(y) \quad (27.187c)$$

$$< \infty. \quad (27.187d)$$

La dernière ligne est l'hypothèse que  $N_1 \in L^1(Y)$ . Le fait que  $\int_X F(x) < \infty$  implique que  $F(x) < \infty$  pour presque tout  $x$ .

(1c) **Pour** (2) Nous avons un calcul comme d'habitude en utilisant Fubini-Tonelli :

$$\|N_2\|_1 = \int_X \left( \int_Y |h(x, y)| dy \right) dx \quad (27.188a)$$

$$= \int_Y \|h_{1,y}\|_1 dy \quad (27.188b)$$

$$= \int_Y N_1(y) dy \quad (27.188c)$$

$$< \infty. \quad (27.188d)$$

(2) Pour  $p = \infty$  Nous définissons

$$\begin{aligned} F: X &\rightarrow \mathbb{C} \\ x &\mapsto \int_Y h(x, y) dy. \end{aligned} \quad (27.189)$$

Nous avons  $|F(x)| \leq \int_Y |h(x, y)| dy$ , et  $|h_{1,y}(x)| \leq \|h_{1,y}\|_\infty$  pour presque tout  $x \in X$ . Armé de cela, nous pouvons faire un petit calcul avec le supremum essentiel (27.78) :

$$\|F\|_{L^\infty} = \text{ess-sup}_{x \in X} \left| \int_Y h(x, y) dy \right| \quad (27.190a)$$

$$\leq \text{ess-sup}_{x \in X} \int_Y \|h_{1,y}\|_{L^\infty} dy \quad (27.190b)$$

$$= \int_Y \|h_{1,y}\|_{L^\infty} dy. \quad (27.190c)$$

Cela est bien la relation (27.183) avec  $p = \infty$ .

(3) Si  $1 < p < \infty$  Plusieurs parties.

(3a) Un premier calcul

Nous posons

$$\begin{aligned} F: X &\rightarrow \mathbb{C} \\ x &\mapsto \int_Y h(x, y) dy, \end{aligned} \quad (27.191)$$

nous considérons un certain  $\varphi \in L^q(X)$ , et nous calculons  $s = \left| \int_X F\varphi \right|$  :

$$s \leq \int_X |F||\varphi| \quad (27.192a)$$

$$= \int_X \left[ \int_Y |h(x, y)| |\varphi(x)| dy \right] dx. \quad (27.192b)$$

Le théorème de Fubini-Tonelli 14.302(2) appliqué à la fonction  $f(x, y) = |h(x, y)| |\varphi(x)|$  permet de permuter les intégrales parce que  $f$  est mesurable et positive<sup>22</sup> :

$$s \leq \int_Y \left( \int_X |h(x, y)| |\varphi(x)| dx \right) dy. \quad (27.193)$$

Nous pouvons utiliser Hölder 27.39(1) pour l'intégrale sur  $X$  parce que  $\varphi \in L^q$  et  $h_{1,y} \in L^p$  :

$$\int_X |h(x, y)| |\varphi(x)| = \|h_{1,y}\varphi\|_1 \leq \|h_{1,y}\|_p \|\varphi\|_q = N_1(y) \|\varphi\|_q. \quad (27.194)$$

Parmi les hypothèses, nous avons  $N_1 \in L^1(Y)$ . Donc l'intégrale sur  $Y$  se passe en réalité assez bien et nous avons

$$s \leq \int_Y \|h_{1,y}\|_p \|\varphi\|_q dy = \|\varphi\|_q \int_Y N_1(y) dy < \infty. \quad (27.195)$$

Nous avons prouvé que pour tout  $\varphi \in L^q(Y)$ ,

$$\left| \int_X F\varphi \right| \leq \|\varphi\|_q \|N_1\|_1 < \infty. \quad (27.196)$$

Nous prenons maintenant le supremum sur tous les  $\varphi \in L^q(Y)$  satisfaisant  $\|\varphi\|_q = 1$ . La proposition 27.47(2) nous dit que le membre de gauche devient  $\|F\|_p$ . Quant au membre de droite, il devient simplement  $\|N_1\|_1$  :

$$\|F\|_p \leq \|N_1\|_1 < \infty. \quad (27.197)$$

EQooKIFGooRHBGXl

22. Nous n'avons pas (encore) de garanties que  $|f(x, y)| < \infty$ , mais ce n'est pas important pour le moment.



- (3b) **Un second calcul** Nous faisons exactement le même calcul que le premier, mais en partant de

$$\begin{aligned} G: X &\rightarrow [0, \infty] \\ x &\mapsto \int_Y |h(x, y)| dy. \end{aligned} \quad (27.198)$$

Nous considérons  $\varphi \in L^q(X)$ , et nous calculons en suivant les mêmes étapes :

$$\|G\varphi\|_1 = \int_X |G(x)\varphi(x)| dx \quad (27.199a)$$

$$= \dots \quad (27.199b)$$

$$\leq \|\varphi\|_q \|N_1\|_1. \quad (27.199c)$$

Cela montre que pour tout  $\varphi \in L^q(X)$  nous avons

$$\left| \int_X G\varphi \right| \leq \int_X |G\varphi| = \|G\varphi\|_1 \leq \|\varphi\|_q \|N_1\|_1 < \infty, \quad (27.200)$$

et en passant au supremum, nous avons

$$\|G\|_p \leq \|N_1\|_1 < \infty \quad (27.201)$$

- (3c) **Pour (3)** Nous avons, en utilisant 27.197 :

$$\left( \int_X \left| \int_Y h(x, y) dy \right|^p dx \right)^{1/p} = \left( \int_X |F(x)|^p dx \right)^{1/p} = \|F\|_p \leq \|N_1\|_1 \quad (27.202)$$

- (3d) **Pour (1)** Nous avons vu que  $\|G\|_p < \infty$ . Donc  $G < \infty$  presque partout. Cela prouve que  $h_{2,x} \in L^1(Y)$  pour presque tout  $x$ .

- (3e) **Pour (2)** Nous devons voir que  $N_2 \in L^p(X)$ . Calcul avec permutation d'intégrales de Fubini-Tonelli :

$$\|N_2\|_p^p = \int_X \left| \|h_{2,x}\|_1 \right|^p dx \quad (27.203a)$$

$$= \int_X \left( \int_Y |f(x, y)| dy \right)^p dx \quad (27.203b)$$

$$= \int_X |G(x)|^p \quad (27.203c)$$

$$= \|G\|_p^p \quad (27.203d)$$

$$< \infty. \quad (27.203e)$$

□

LEMooPSBWooGLggTe

**Lemme 27.51** ([724]).

Si  $(X, \mu)$  est un espace mesuré  $\sigma$ -fini<sup>23</sup> non nul<sup>24</sup>, il existe une fonction mesurable strictement positive en escalier  $k: X \rightarrow ]0, \infty[$  telle que

$$\int_X k d\mu = 1. \quad (27.204)$$

*Démonstration.* Vue que  $X$  est  $\sigma$ -fini, il existe des parties  $(A_n)$  de mesures finies telles que  $X = \bigcup_{i=1}^{\infty} A_n$ . Afin d'avoir des parties disjointes, nous posons

$$B_n = A_n \setminus \left( \bigcup_{i=1}^{n-1} A_i \right). \quad (27.205)$$

23. Définition 14.18.

24. C'est-à-dire que  $\mu$  n'est pas la mesure identiquement nulle.

Nous posons

$$k = \sum_{n=0}^{\infty} \frac{\alpha_n}{\mu(B_n)} \mathbb{1}_{B_n} \quad (27.206)$$

où  $(\alpha_n)$  est une suite de réels strictement positifs dont la somme vaut 1. En vertu de la somme géométrique (10.404), prendre  $\alpha_n = 2^{-n}$  fonctionne.

Notons que les  $B_n$  étant disjoints, pour chaque  $x$ , la somme définissant  $k(x)$  est non seulement finie, mais se réduit à un seul terme. L'application  $k$  est en escalier ; son intégrale vaut

$$\int_X k d\mu = \sum_{n=0}^{\infty} \frac{\alpha_n}{\mu(B_n)} \mu(B_n) = \sum_{n=1}^{\infty} \alpha_n = 1. \quad (27.207)$$

□

À mon avis l'énoncé de 27.52 n'est pas tellement correct. Il manque des hypothèses. Ne l'utilisez que si vous êtes sûre ; et préférez 27.50.

PROPOoGZJZooXfZdqn

**Proposition 27.52** (Forme intégrale de Minkowski[724]).

Soient deux espaces  $\sigma$ -finis  $(X, \mu)$  et  $(Y, \nu)$ . Si  $f: X \times Y \rightarrow \mathbb{C}$  est mesurable<sup>25</sup> et si  $p \geq 1$ , alors

$$\left[ \int_X \left( \int_Y |f(x, y)| d\nu(y) \right)^p d\mu(x) \right]^{1/p} \leq \int_Y \left( \int_X |f(x, y)|^p d\mu(x) \right)^{1/p} d\nu(y). \quad (27.208)$$

EQooAEXWooYJtGGB

De façon plus compacte :

$$\left\| x \mapsto \int_Y f(x, y) d\nu(y) \right\|_p \leq \int_Y \|f_y\|_p d\nu(y) \quad (27.209)$$

où  $f_y(x) = f(x, y)$ .

*Démonstration.* En plusieurs parties.

(1) **La fonction  $M$**  Pour toute application mesurable  $g: X \times Y \rightarrow \mathbb{C}$  nous posons

$$M(g) = \int_Y \left[ \int_X |g(x, y)|^p d\mu(x) \right]^{1/p} d\nu(y). \quad (27.210)$$

Cette fonction a une propriété d'homogénéité :  $M(\lambda g) = \lambda M(g)$  pour tout  $\lambda \in [0, \infty[$ .

(2) **Si  $M(g) < 1$**  Nous considérons une fonction mesurable  $g: X \times Y \rightarrow \mathbb{C}$  telle que  $M(g) < 1$ . Soit  $k: Y \rightarrow ]0, \infty[$  une fonction donnée par le lemme 27.51, c'est-à-dire telle que  $\int_Y k = 1$ . Nous posons

$$s: Y \rightarrow [0, \infty]$$

$$y \mapsto \left( \int_X |g(x, y)|^p d\mu(x) \right)^{1/p}, \quad (27.211)$$

ainsi que

$$h(y) = s(y) + (1 - M(g))k(y). \quad (27.212)$$

Vu que  $M(g) < 1$  et que  $k > 0$ , nous avons l'inégalité stricte  $h(y) > s(y) \geq 0$  pour tout  $y \in Y$ . Nous avons de plus

$$\int_Y h(y) d\nu(y) = \underbrace{\int_Y s}_{=M(g)} + (1 - M(g)) \underbrace{\int_Y k}_{=1} = 1. \quad (27.213)$$

25. Pour l'espace mesuré produit  $X \times Y, \mu \otimes \nu$ .

Nous considérons maintenant la mesure  $\rho = h\nu$  (produite d'une mesure par une fonction, définition 14.212). Nous venons de montrer que c'est une mesure de probabilité. En posant  $\alpha(x) = |x|^p$ , nous pouvons faire quelques calculs avec les justifications en-dessous :

$$\alpha\left(\int_Y |g(x, y)| d\nu(y)\right) = \alpha\left(\int_Y h(y)^{-1} |g(x, y)| h(y) d\nu(y)\right) \quad (27.214a)$$

$$= \alpha\left(\int_Y [|g(x, y)| h(y)^{-1}] d\rho(y)\right) \quad (27.214b)$$

$$\leq \int_Y \alpha(|g(x, y)| h(y)^{-1}) d\rho(y) \quad (27.214c)$$

$$= \int_Y |g(x, y)|^p |h(y)|^{1-p} d\nu(y) \quad (27.214d)$$

$$(27.214e)$$

Justifications.

- Pour (27.214c). Inégalité de Jensen, proposition 27.37
- Pour (27.214d). La fonction  $h$  est à valeurs positives ; donc  $h = |h|$  et nous pouvons coller le  $h$  qui ressort de  $\rho$  avec le  $|h^{-1}|^p$ .

Nous continuons en intégrant l'inégalité (27.214) sur  $X$  :

$$\int_X \left(\int_Y |g(x, y)| d\nu(y)\right)^p d\mu(x) \leq \int_X \left[\int_Y |g(x, y)|^p h(y)^{1-p} d\nu(y)\right] d\mu(x) \quad (27.215a)$$

$$= \int_Y \int_X [|g(x, y)|^p h(y)^{1-p} d\mu(x)] d\nu(y) \quad (27.215b)$$

$$= \int_Y \left(\int_X [|g(x, y)|^p d\mu(x)]\right) h(y)^{1-p} d\nu(y) \quad (27.215c)$$

$$= \int_Y s(y)^p h(y)^{1-p} d\nu(y) \quad (27.215d)$$

$$\leq \int_Y h(y) d\nu(y) \quad (27.215e)$$

$$= 1. \quad (27.215f)$$

Justifications.

- Pour (27.215b). Les intégrales sont permutées grâce au théorème de Fubini-Tonelli 14.302.
- Pour (27.215e). Nous savons que  $h > s \geq 0$ , donc  $s(y)^p/h(y)^p < 1$ .

Au final nous avons prouvé que si  $M(g) < 1$ , alors nous avons

$$\int_X \left(\int_Y |g(x, y)| d\nu(y)\right)^p d\mu(x) \leq 1. \quad (27.216)$$

- (3) **La preuve proprement dite** Nous considérons une fonction mesurable  $f: X \times Y \rightarrow \mathbb{C}$  comme dans les hypothèses. Astuce : considérons  $0 < t < 1$  et appliquons l'inégalité (27.216) à  $g = \frac{t}{M(f)} f$ . Cela est possible parce que

$$M(g) = \frac{t}{M(f)} M(f) = t < 1. \quad (27.217)$$

Nous avons :

$$\int_X \left(\int_Y \left|\frac{t}{M(f)} f(x, y)\right|^p d\nu(y)\right)^p d\mu(x) \leq 1, \quad (27.218)$$

et donc, vu que  $M(f)$  et  $t$  sont positifs,

$$\int_X \left(\int_Y |f(x, y)|^p d\nu(y)\right)^p d\mu(x) \leq \frac{M(f)^p}{t^p}. \quad (27.219)$$

Cette inégalité est valable pour tout  $t < 1$  et, ô incroyable!, ce  $t$  n'est qu'à droite. Nous pouvons prendre la limite  $t \rightarrow 1$  dans cette inégalité<sup>26</sup> :

$$\int_X \left( \int_Y |f(x, y)|^p d\nu(y) \right)^p d\mu(x) \leq M(f). \quad (27.220)$$

Et ça, c'est ce que nous devons démontrer.

□

### 27.3.7 Ni inclusions ni inégalités

Aucun espace  $L^p(\mathbb{R})$  n'est inclus dans aucun autre ni aucune norme n'est plus grande qu'une autre (sur les intersections). Nous verrons cependant en la proposition 27.53 que de telles inclusions et inégalités sont possibles pour  $L^p([0, 1])$ .

Nous allons donner des exemples de tout ça en supposant  $p < q$  et en nous appuyant lourdement sur les intégrales de  $\frac{1}{x^\alpha}$  étudiées par la proposition 14.301.

(1)  $L^p \not\subseteq L^q$

La fonction

$$f(x) = \begin{cases} \frac{1}{x^{1/q}} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases} \quad (27.221)$$

est dans  $L^p$  mais pas dans  $L^q$ . En effet

$$\|f\|_p^p = \int_0^1 \frac{1}{x^{p/q}} dx < \infty \quad (27.222)$$

parce que  $p < q$  et  $p/q < 1$ . Par contre

$$\|f\|_q^q = \int_0^1 \frac{1}{x} dx = \infty. \quad (27.223)$$

(2)  $L^q \not\subseteq L^p$

La fonction

$$f(x) = \begin{cases} \frac{1}{x^{1/p}} & \text{si } x > 1 \\ 0 & \text{sinon} \end{cases} \quad (27.224)$$

est dans  $L^q$  mais pas dans  $L^p$ . En effet

$$\|f\|_p^p = \int_1^\infty \frac{1}{x} = \infty \quad (27.225)$$

alors que

$$\|f\|_q^q = \int_1^\infty \frac{1}{x^{q/p}} dx < \infty. \quad (27.226)$$

(3) Exemple de  $\|f\|_p > \|f\|_q$

La fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in [0, 2] \\ 0 & \text{sinon.} \end{cases} \quad (27.227)$$

Nous avons

$$\|f\|_p = 2^{1/p} \quad (27.228a)$$

$$\|f\|_q = 2^{1/q}. \quad (27.228b)$$

Mais comme  $p < q$  donc  $\|f\|_p > \|f\|_q$ .

26. Nous ne pouvons pas simplement poursuivre les inégalités en majorant  $t$  par 1 parce que le  $t$  est au dénominateur : l'inégalité irait dans le mauvais sens.

(4) **Exemple de  $\|f\|_p < \|f\|_q$** 

La fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}] \\ 0 & \text{sinon.} \end{cases} \quad (27.229)$$

Alors

$$\|f\|_p = \frac{1}{2^{1/p}} \quad (27.230a)$$

$$\|f\|_q = \frac{1}{2^{1/q}} \quad (27.230b)$$

et donc  $\|f\|_p < \|f\|_q$ .

Ces exemples donnent un exemple de fonction  $f$  telle que  $\|f\|_p < \|f\|_q$  pour tout espace  $L^p(I)$  et  $L^q(I)$  avec  $I \subset \mathbb{R}$ . Par contre l'exemple  $\|f\|_p > \|f\|_q$  ne fonctionne que si la taille de  $I$  est plus grande que 1. Et pour cause : il y a des inclusions si  $I$  est borné.

PropIRDooFSWORl

**Proposition 27.53** ([13]).

Inclusions et inégalités dans le cas d'un ensemble de mesure finie.

(1) Soit  $(\Omega, \mathcal{A}, \mu)$  un espace mesuré fini et  $1 \leq p \leq +\infty$ . Alors  $L^q(\Omega) \subset L^p(\Omega)$  dès que  $p \leq q$ .

ItemWSTooLcpOvXii

(2) Si  $1 < p < 2$  et si  $f \in L^2([0, 1])$  alors  $\|f\|_p \leq \|f\|_2$ .

*Démonstration.* Pour la simplicité des notations nous allons noter  $L^p$  pour  $L^p(\Omega)$ , et pareillement pour  $L^q$ . Soit  $f \in L^q$ . Nous posons

$$A = \{x \in \Omega \text{ tel que } |f(x)| \geq 1\}. \quad (27.231)$$

Étant donné que  $p \leq q$  nous avons  $|f|^p \leq |f|^q$  sur  $A$ ; par conséquent  $\int_A |f|^p$  converge parce que  $\int_A |f|^q$  converge.

L'ensemble  $A^c$  est évidemment borné (complémentaire dans  $\Omega$ ) et sur  $A^c$  nous avons  $|f(x)| \leq 1$  et donc  $|f|^p \leq 1$ . L'intégrale  $\int_{A^c} |f|^p$  converge donc également.

Au final  $\int_\Omega |f|^p$  converge et  $f \in L^p$ .

Soit à présent  $f \in L^2$ ; par le premier point nous avons immédiatement  $f \in L^2 \cap L^p$ . Soit aussi  $r \in \mathbb{R}$  tel que  $\frac{1}{2/p} + \frac{1}{r} = 1$ . Nous avons  $|f|^p \in L^{2/p}$ , et vu que nous sommes sur un domaine borné,  $1 \in L^r$ . Nous écrivons l'inégalité de Hölder (27.102) avec ces fonctions. D'une part

$$\|f\|_1 = \||f|^p\|_1 = \|f\|_p^p. \quad (27.232)$$

D'autre part

$$\||f|^p\|_{2/p} = \left( \int |f|^2 \right)^{p/2} = \|f\|_2^p. \quad (27.233)$$

Donc  $\|f\|_p^p \leq \|f\|_2^p$ , ce qui prouve l'assertion (2) parce que  $p > 1$ . □

**Remarque 27.54.**

Nous n'avons cependant pas  $L^2([0, 1]) = L^p([0, 1])$  parce que l'exemple (27.221) fonctionne encore :

$$f(x) = \frac{1}{\sqrt{x}} \quad (27.234)$$

pour  $x \in [0, 1]$  donne bien

$$\|f\|_2 = \int_0^1 \frac{1}{x} = \infty \quad (27.235)$$

et  $\|f\|_p = \int_0^1 \frac{1}{x^{p/2}} < \infty$  parce que  $1 < p < 2$ .

### 27.3.8 Complétude

**Théorème 27.55** ([733, 734]).

Pour  $1 \leq p < \infty$ , l'espace  $L^p(\Omega, \mathcal{A}, \mu)$  est complet.

*Démonstration.* Soit  $(f_n)_{n \in \mathbb{N}}$  une suite de Cauchy dans  $L^p$ . Pour tout  $i$ , il existe  $N_i \in \mathbb{N}$  tel que  $\|f_k - f_l\|_p \leq 2^{-i}$  pour tout  $k, l \geq N_i$ . Nous considérons la sous-suite  $g_i = f_{N_i}$ , de telle sorte qu'en particulier

$$\|g_i - g_{i-1}\|_p \leq 2^{-i+1}. \quad (27.236)$$

Pour chaque  $j$  nous considérons la somme télescopique

$$g_j = g_0 + \sum_{i=1}^j (g_i - g_{i-1}) \quad (27.237)$$

et l'inégalité

$$|g_j| \leq |g_0| + \sum_{i=1}^j |g_i - g_{i-1}|. \quad (27.238)$$

Nous allons noter

$$h_j = |g_0| + \sum_{i=1}^j |g_i - g_{i-1}|. \quad (27.239)$$

La suite de fonctions  $(h_j)$  ainsi définie est une suite croissante de fonctions positives qui converge donc (ponctuellement) vers une fonction  $h$  qui peut éventuellement valoir l'infini en certains points. Par continuité de la fonction  $x \mapsto x^p$  nous avons

$$\lim_{j \rightarrow \infty} h_j^p = h^p, \quad (27.240)$$

puis par le théorème de la convergence monotone (théorème 14.178) nous avons

$$\lim_{j \rightarrow \infty} \int_{\Omega} h_j^p d\mu = \int_{\Omega} h^p d\mu. \quad (27.241)$$

Utilisant à présent la continuité de la fonction  $x \mapsto x^{1/p}$  nous trouvons

$$\lim_{j \rightarrow \infty} \left( \int h_j^p \right)^{1/p} = \left( \int |h|^p \right)^{1/p}. \quad (27.242)$$

Nous avons donc déjà montré que

$$\lim_{j \rightarrow \infty} \|h_j\|_p = \left( \int |h|^p \right)^{1/p} \quad (27.243)$$

où, encore une fois, rien ne garantit à ce stade que l'intégrale à droite soit un nombre fini. En utilisant l'inégalité de Minkowski (proposition 27.45) et l'inégalité (27.236) nous trouvons

$$\|h_j\|_p \leq \|g_0\|_p + \sum_{i=1}^j \|g_i - g_{i-1}\|_p \leq \|g_0\|_p + 1. \quad (27.244)$$

En passant à la limite,

$$\left( \int |h|^p \right)^{1/p} = \lim_{j \rightarrow \infty} \|h_j\|_p \leq \|g_0\|_p + 1 < \infty. \quad (27.245)$$

Par conséquent  $\int |h|^p$  est finie et

$$h \in L^p(\Omega, \mathcal{A}, \mu). \quad (27.246)$$

En particulier, l'intégrale  $\int h$  est finie (parce que  $p \geq 1$ ) et donc que  $h(x) < \infty$  pour presque tout  $x \in \Omega$ .

Nous savons que  $h(x)$  est la limite des sommes partielles (27.239), en particulier la série

$$\sum_{j=1}^{\infty} |g_j - g_{j-1}| \quad (27.247)$$

converge ponctuellement. En vertu du corolaire 14.209, la série de terme général  $g_j - g_{j-1}$  converge ponctuellement. La suite  $g_j$  converge donc vers une fonction que nous notons  $g$ . Par ailleurs la suite  $g_j$  est dominée par  $h \in L^p$ , le théorème de la convergence dominée (théorème 14.207) implique que

$$\lim_{j \rightarrow \infty} \|g_j - g\|_p = 0. \quad (27.248)$$

Nous allons maintenant prouver que  $\lim_{n \rightarrow \infty} \|f_n - g\|_p = 0$ . Soit  $\epsilon > 0$ . Pour tout  $n$  et  $i$  nous avons

$$\|f_n - g\|_p = \|f_n - f_{N_i} + f_{N_i} - g\|_p \leq \|f_n - f_{N_i}\|_p + \|f_{N_i} - g\|_p. \quad (27.249)$$

Pour rappel,  $f_{N_i} = g_i$ . Si  $i$  et  $n$  sont suffisamment grands nous pouvons obtenir que chacun des deux termes est plus petit que  $\epsilon/2$ .

Il nous reste à prouver que  $g \in L^p(\Omega, \mathcal{A}, \mu)$ . Nous avons déjà vu (équation (27.246)) que  $h \in L^p$ , mais  $|g_i| \leq h^p$ , par conséquent  $g \in L^p$ .

Nous avons donc montré que la suite de Cauchy  $(f_n)$  converge vers une fonction de  $L^p$ , ce qui signifie que  $L^p$  est complet.  $\square$

ThoGVmqOro

**Théorème 27.56** (Riesz-Fischer[126, 735]).

Soient un espace mesuré  $(\Omega, \mathcal{A}, \mu)$  et  $p \in [1, \infty]$ . Alors

ItemPDnjOJzi

(1) Toute suite convergente dans  $L^p(\Omega, \mathcal{A}, \mu)$  admet une sous-suite convergente presque partout sur  $\Omega$ .

ItemPDnjOJzii

(2) La sous-suite donnée en (1) est dominée par un élément de  $L^p(\Omega, \mathcal{A}, \mu)$ .

ItemPDnjOJzii

(3) L'espace  $L^p(\Omega, \mathcal{A}, \mu)$  est de Banach.

*Démonstration.* Le cas  $p = \infty$  est à séparer des autres valeurs de  $p$  parce qu'on y parle de norme uniforme, et aucune sous-suite n'est à considérer.

(1) **Cas  $p = \infty$ .** Nous commençons par prouver dans le cas  $p = \infty$ . Soit  $(f_n)$  une suite de Cauchy dans  $L^\infty(\Omega)$ , ou plus précisément une suite de représentants d'éléments de  $L^p$ . Pour tout  $k \geq 1$ , il existe  $N_k \geq 0$  tel que si  $m, n \geq N_k$ , on a

$$\|f_m - f_n\|_\infty \leq \frac{1}{k}. \quad (27.250)$$

En particulier, l'ensemble  $E_k$  sur lequel

$$|f_m(x) - f_n(x)| \geq \frac{1}{k} \quad (27.251)$$

est de mesure nulle. D'après le lemme 14.28, la partie  $E = \bigcup_{k \in \mathbb{N}} E_k$ , est encore de mesure nulle. En résumé, nous avons un  $N_k$  tel que si  $m, n \geq N_k$ , alors

$$|f_n(x) - f_m(x)| \leq \frac{1}{k} \quad (27.252)$$

EqKAWSpmtG

pour tout  $x$  hors de  $E$ . Donc pour chaque  $x \in \Omega \setminus E$ , la suite  $n \mapsto f_n(x)$  est de Cauchy dans  $\mathbb{R}$  et converge donc. Cela définit donc une fonction

$$\begin{aligned} f: \Omega \setminus E &\rightarrow \mathbb{R} \\ x &\mapsto \lim_{n \rightarrow \infty} f_n(x). \end{aligned} \quad (27.253)$$

Cela prouve le point (1) : la convergence ponctuelle.

En passant à la limite  $n \rightarrow \infty$  dans l'équation (27.252) et tenant compte que cette majoration tient pour presque tout  $x$  dans  $\Omega$ , nous trouvons

$$\|f - f_n\|_\infty \leq \frac{1}{k}. \quad (27.254)$$

Donc non seulement  $f$  est dans  $L^\infty$ , mais en plus la suite  $(f_n)$  converge vers  $f$  au sens  $L^\infty$ , c'est-à-dire uniformément. Cela prouve le point (3). En ce qui concerne le point (2), la suite  $f_n$  est entièrement (à partir d'un certain point) dominée par la fonction  $1 + |f|$  qui est dans  $L^\infty$ .

(2) Cas  $p < \infty$ .

Toute suite convergente étant de Cauchy, nous considérons une suite de Cauchy  $(f_n)$  dans  $L^p(\Omega)$  et ce sera suffisant pour travailler sur le premier point. Pour montrer qu'une suite de Cauchy converge, il est suffisant de montrer qu'une sous-suite converge. Soit  $\varphi: \mathbb{N} \rightarrow \mathbb{N}$  une fonction strictement croissante telle que pour tout  $n \geq 1$  nous ayons

$$\|f_{\varphi(n+1)} - f_{\varphi(n)}\|_p \leq \frac{1}{2^n}. \quad (27.255)$$

Pour créer la fonction  $\varphi$ , il est suffisant de prendre le  $N_k$  donné par la condition de Cauchy pour  $\epsilon = 1/2^k$  et de considérer la fonction définie par récurrence<sup>27</sup> par  $\varphi(1) = N_1$  et  $\varphi(n+1) > \max\{N_n, \varphi(n)\}$ . Ensuite nous considérons la fonction

$$g_n(x) = \sum_{k=1}^n |f_{\varphi(k+1)}(x) - f_{\varphi(k)}(x)|. \quad (27.256)$$

Notons que pour écrire cela nous avons considéré des représentants  $f_k$  qui sont alors des fonctions à l'ancienne. Étant donné que  $g_n$  est une somme de fonctions dans  $L^p$ , c'est une fonction  $L^p$ , comme nous pouvons le constater en calculant sa norme :

$$\|g_n\|_p \leq \sum_{k=1}^n \|f_{\varphi(k+1)} - f_{\varphi(k)}\|_p \leq \sum_{k=1}^n \frac{1}{2^k} \leq \sum_{k=1}^{\infty} \frac{1}{2^k} = 1. \quad (27.257)$$

Étant donné que tous les termes de la somme définissant  $g_n$  sont positifs, la suite  $(g_n)$  est croissante. Mais elle est bornée en norme  $L^p$  et donc sujette à obéir au théorème de Beppo-Levi 14.178 sur la convergence monotone. Il existe donc une fonction  $g \in L^p(\Omega)$  telle que  $g_n \rightarrow g$  presque partout.

Soit un  $x \in \Omega$  pour lequel  $g_n(x) \rightarrow g(x)$  ; alors pour tout  $n \geq 2$  et  $\forall q \geq 0$ ,<sup>E<sub>q</sub>WTHo<sub>j</sub>C<sub>q</sub></sup>

$$|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| = \left| \underbrace{\varphi(n+q)(x) + \sum_{k=1}^{q-1} f_{\varphi(n+k)}(x)}_{=\sum_{k=1}^q} \underbrace{\sum_{k=1}^{q-1} f_{\varphi(n+k)}(x) - f_{\varphi(n)}(x)}_{=\sum_{k=1}^q} \right| \quad (27.258a)$$

$$= \left| \sum_{k=1}^q f_{\varphi(n+k)}(x) - \sum_{k=1}^q f_{\varphi(n+k-1)}(x) \right| \quad (27.258b)$$

$$\leq \sum_{k=1}^q |f_{\varphi(n+k)}(x) - f_{\varphi(n+k-1)}(x)| \quad (27.258c)$$

$$= g_{n+q+1}(x) - g_{n+1}(x) \quad (27.258d)$$

$$\leq g(x) - g_{n+1}(x). \quad (27.258e)$$

Nous prenons la limite  $n \rightarrow \infty$  ; la dernière expression tend vers zéro et donc

$$|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| \rightarrow 0 \quad (27.259)$$

27. Utilisation du théorème 1.120. Vous n'êtes pas obligé de le citer à chaque fois, mais c'est bien de garder en tête que la définition de fonctions par récurrence n'est pas quelque chose de complètement trivial.



pour tout  $q$ . Donc pour presque tout  $x \in \Omega$ , la suite  $n \mapsto f_{\varphi(n)}(x)$  est de Cauchy dans  $\mathbb{R}$  et donc y converge vers un nombre que nous nommons  $f(x)$ . Cela définit une fonction

$$\begin{aligned} f: \Omega \setminus E &\rightarrow \mathbb{R} \\ x &\mapsto \lim_{n \rightarrow \infty} f_{\varphi(n)}(x) \end{aligned} \quad (27.260)$$

où  $E$  est de mesure nulle. Montrons que  $f$  est bien dans  $L^p(\Omega)$ ; pour cela nous complétons la série d'inégalités (27.258) en

$$|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| \leq g(x) - g_{n-1}(x) \leq g(x). \quad (27.261)$$

En prenant la limite  $q \rightarrow \infty$  nous avons l'inégalité

$$|f(x) - f_{\varphi(n)}(x)| \leq g(x) \quad (27.262)$$

pour presque tout  $x \in \Omega$ , c'est-à-dire pour tout  $x \in \Omega \setminus E$ . Cette inégalité implique deux choses valables pour presque tout  $x$  dans  $\Omega$  :

$$f(x) \in B(g(x), f_{\varphi(n)}(x)) \quad (27.263a)$$

$$f_{\varphi(n)}(x) \leq |f(x)| + |g(x)|. \quad (27.263b)$$

La première inégalité assure que  $|f|^p$  est intégrable sur  $\Omega \setminus E$  parce que  $|f|$  est majorée par  $|g| + |f_{\varphi(n)}|$ . Elle prouve par conséquent le point (1) parce que  $n \mapsto f_{\varphi(n)}$  est une sous-suite convergente presque partout. La seconde montre le point (2).

Attention : à ce point nous avons prouvé que  $n \mapsto f_{\varphi(n)}$  est une suite de fonctions qui converge *ponctuellement presque partout* vers une fonction  $f$  qui s'avère être dans  $L^p$ . Nous n'avons pas montré que cette suite convergeait au sens de  $L^p$  vers  $f$ . Ce que nous devons montrer est que

$$\|f - f_{\varphi(n)}\|_p \rightarrow 0. \quad (27.264)$$

L'inégalité (27.262) nous donne aussi, toujours pour presque tout  $x \in \Omega$  :

$$|f(x) - f_{\varphi(n)}(x)|^p \leq g(x)^p \quad (27.265)$$

ce qui signifie que la suite  $|f - f_{\varphi(n)}|^p$  est dominée par la fonction  $|g|^p$  qui est intégrable sur  $\Omega \setminus E$  et tout autant sur  $\Omega$  parce que  $E$  est négligeable; cela prouve au passage le point (2), et le théorème de la convergence dominée de Lebesgue (14.207) nous dit que

$$\lim_{n \rightarrow \infty} \int_{\Omega} |f(x) - f_{\varphi(n)}(x)|^p dx = \int_{\Omega} \lim_{n \rightarrow \infty} |f(x) - f_{\varphi(n)}(x)|^p dx = 0. \quad (27.266)$$

Cette dernière suite d'égalités se lit de la façon suivante :

$$\lim_{n \rightarrow \infty} \|f - f_{\varphi(n)}\|_p = \left\| \lim_{n \rightarrow \infty} |f - f_{\varphi(n)}| \right\|_p = 0. \quad (27.267)$$

Nous en déduisons que la suite  $n \mapsto f_{\varphi(n)}$  est convergente vers  $f$  au sens de la norme  $L^p(\Omega)$ . Or la suite de départ  $(f_n)$  était de Cauchy (pour la norme  $L^p$ ); donc l'existence d'une sous-suite convergente implique la convergence de la suite entière vers  $f$ , ce qu'il fallait démontrer.  $\square$

Le théorème suivant est souvent cité en disant que  $L^p$  est un espace de Hilbert si et seulement si  $p = 2$ . Comme vous le voyez, il faut un peu plus d'hypothèses.

Je précise que je suis le seul à nommer ce théorème par le nom de Weinersmith. Je ne sais pas si il a déjà un nom; alors pourquoi pas celui-là plutôt qu'un autre? La raison de ce choix est dans la constante de Wiener, définition 44.1.

28. À ce point, [126] se contente de majorer  $|f_{\varphi(n)}(x)|$  par  $|f(x)| + |g(x)|$ , mais je ne comprends pas comment cette majoration nous permet d'utiliser la convergence dominée de Lebesgue pour montrer (27.264).

THOooCCMBooGulxkQ

**Théorème 27.57** (Théorème de Weinersmith[736, 737]).

Nous considérons un espace mesuré  $(\Omega, \mathcal{A}, \mu)$  ainsi qu'un nombre  $p \in [1, \infty]$ . Nous supposons

- (1)  $L^p(\Omega, \mathcal{A}, \mu)$  est un espace de Hilbert,
- (2) Il existe des parties  $A, B \subset \Omega$  telles que  $A \cap B = \emptyset$  et  $0 < \mu(A) < \infty$  et  $0 < \mu(B) < \infty$ .

Alors  $p = 2$ .

*Démonstration.* Vu que  $L^p$  est un espace de Hilbert (hypothèse), il vérifie l'identité du parallélogramme de la proposition 10.27, c'est-à-dire

$$\|f + g\|_p^2 + \|f - g\|_p^2 = 2\|f\|_p^2 + 2\|g\|_p^2. \quad (27.268)$$

EQooAKKYooURlhwj

- (1) **Pour**  $1 \leq p < \infty$

Soient donc  $A, B$  comme dans l'hypothèse. Nous considérons les fonctions

$$f = \frac{1}{\mu(A)^{1/p}} \mathbb{1}_A \quad (27.269a)$$

$$g = \frac{1}{\mu(B)^{1/p}} \mathbb{1}_B. \quad (27.269b)$$

En ce qui concerne les normes  $L^p$  de  $f$  et  $g$ , c'est un calcul simple :

$$\|f\|_p^2 = \left( \int_{\Omega} |f(\omega)|^p d\mu \right)^{2/p} = \left( \int_A \frac{1}{\mu(A)} d\mu \right)^{2/p} = 1. \quad (27.270)$$

De même pour  $g$  :  $\|f\|_p^2 = \|g\|_p^2 = 1$ . Donc

$$2\|f\|_p^2 + 2\|g\|_p^2 = 4 \quad (27.271)$$

En ce qui concerne la somme,

$$\|f + g\|_p^2 = \left( \int_A \frac{d\mu}{\mu(A)} + \int_B \frac{d\mu}{\mu(B)} \right)^{2/p} = 2^{2/p}. \quad (27.272)$$

Pour la différence, la seule subtilité à voir est que

$$\int_{\Omega} |\mathbb{1}_A - \mathbb{1}_B|^p = \int_A |\mathbb{1}_A - \mathbb{1}_B|^p + \int_B |\mathbb{1}_A - \mathbb{1}_B|^p = \int_A |\mathbb{1}_A|^p + \int_B |-\mathbb{1}_B|^p = \int_A \mathbb{1}_A + \int_B \mathbb{1}_B. \quad (27.273)$$

Ce n'est pas de la magie que le moins se change en plus. Bref, pour la différence nous avons

$$\|f - g\|^2 = \left( \int_{\Omega} |f(\omega) - g(\omega)|^p d\mu \right)^{2/p} \quad (27.274a)$$

$$= \left( \int_{\Omega} \left| \frac{1}{\mu(A)^{1/p}} \mathbb{1}_A(\omega) - \frac{1}{\mu(B)^{1/p}} \mathbb{1}_B(\omega) \right|^p d\mu \right)^{2/p} \quad (27.274b)$$

$$= \left( \int_A \left| \frac{\mathbb{1}_A(\omega)}{\mu(A)^{1/p}} \right|^p + \int_B \left| -\frac{\mathbb{1}_B(\omega)}{\mu(B)^{1/p}} \right|^p \right)^{2/p} \quad (27.274c)$$

$$= 2^{2/p}. \quad (27.274d)$$

Donc  $\|f + g\|_p^2 + \|f - g\|_p^2 = 2 \times 2^{2/p}$ .

Vu que  $L^p$  est un espace de Hilbert, nous avons finalement

$$4 = 2 \times 2^{2/p}. \quad (27.275)$$

Cela est uniquement valable pour  $p = 2$ .

- (2) **Pour**  $p = \infty$  Il suffit de prendre  $f = \mathbb{1}_A$  et  $g = \mathbb{1}_B$ . Nous avons alors  $\|f\|_{L^\infty}^2 = \|g\|_{L^\infty}^2 = \|f + g\|_{L^\infty}^2 = \|f - g\|_{L^\infty}^2 = 1$ .

L'égalité (27.268) devient  $2 = 4$ , ce qui est faux.

□

Si  $(\Omega, \mathcal{A}, \mu)$  est un espace mesuré, est-ce que  $L^p(\Omega, \mathcal{A}, \mu)$  est assuré de n'être pas de Hilbert ? Non.

**Exemple 27.58** ([737]).

Soit la mesure de Dirac sur  $\mathbb{R}$ , c'est-à-dire

$$\delta(A) = \begin{cases} 1 & \text{si } 0 \in A \\ 0 & \text{sinon.} \end{cases} \quad (27.276)$$

Nous allons prouver que pour tout  $p \in [0, \infty]$ , l'espace  $L^p(\mathbb{R}, \mathcal{B}(\mathbb{R}), \delta)$  est un espace de Hilbert. Pour cela nous introduisons le produit hermitien

$$\langle f, g \rangle = f(0)\overline{g(0)}. \quad (27.277)$$

Nous avons par ailleurs la norme

$$\|f\|^p = \int_{\mathbb{R}} |f(x)|^p d\delta = f(0)^p. \quad (27.278)$$

Donc oui,  $\|f\|_p = \sqrt{\langle f, f \rangle}$ .

△

### 27.3.9 Théorèmes d'approximation

PROPOUQUBooAWgNhm

**Proposition 27.59** ([507], thème 24).

Soient  $1 \leq p \leq \infty$  et un espace mesuré  $(\Omega, \mathcal{A}, \mu)$ . Alors les fonctions étagées<sup>29</sup> dans  $L^p(\Omega)$  sont denses dans  $L^p(\Omega)$ .

*Démonstration.* Soit  $f \in L^1(\Omega, \mathcal{A}, \mu)$ . Nous supposons dans un premier temps que  $f: \Omega \rightarrow [0, \infty[$ . Pour les fonctions à valeurs dans  $\mathbb{C}$ , nous verrons plus bas.

Notons que la partie  $\{x \in \Omega \text{ tel que } f(x) = \infty\}$  est de mesure nulle, donc nous pouvons vraiment choisir un représentant à valeurs dans  $[0, \infty[$  et non à valeurs dans  $[0, \infty]$  comme le serait un représentant un peu quelconque.

Par le théorème 14.118, il existe une suite croissante de fonctions étagées  $\phi_n: \Omega \rightarrow [0, \infty[$  telles que  $\phi_n \rightarrow f$  ponctuellement. Notons que ce théorème fonctionne parce que les fonctions  $L^p$  (en tout cas leurs représentants) sont mesurables parce que c'est dans la définition 27.12. Notre devoir est maintenant de prouver que sous l'hypothèse que  $f$  est dans  $L^p$ , alors la convergence  $\phi_n \rightarrow f$  est une convergence dans  $L^p$ .

- (1)  $1 \leq p < \infty$  Vu que  $f$  et  $\phi_n$  sont à valeurs positives nous avons  $|f - \phi_n|^p \leq |f|^p$ . Mais par hypothèse  $|f|^p \in L^1(\Omega)$ . Donc la suite  $g_n = |f - \phi_n|^p$  est majorée (uniformément en  $n$ ) par  $|f|^p$  qui est dans  $L^1$ . Le théorème de la convergence dominée permet de permuter

$$\lim_{n \rightarrow \infty} \int g_n = \int \lim_{n \rightarrow \infty} g_n = 0. \quad (27.279)$$

Cela revient à dire que

$$\lim_{n \rightarrow \infty} \|f - \phi_n\|_p^p = \lim_{n \rightarrow \infty} \int |f - \phi_n|^p = \int \lim_{n \rightarrow \infty} |f - \phi_n|^p = 0, \quad (27.280)$$

ce qui signifie que  $\phi_n \xrightarrow{L^p(\Omega)} f$ .

<sup>29</sup>. Définition 14.112. Pour rappel, une fonction est simple lorsqu'elle prend un nombre fini de valeurs, et elle est étagée lorsqu'elle est en outre mesurable.

- (2) **Pour**  $p = \infty$  Si  $f \in L^\infty(\Omega)$ , alors nous pouvons prendre un représentant borné. Avec lui, nous avons  $\phi_n \xrightarrow{\|\cdot\|_\infty} f$ . Avec cela nous avons, pour  $\epsilon$  donné, un  $n$  assez grand pour avoir

$$N_\infty(\phi_n - f) \leq \|\phi_n - f\|_\infty < \epsilon. \quad (27.281)$$

Si  $f$  est à valeurs dans  $\mathbb{C}$  au lieu de  $[0, \infty[$ , il suffit de faire valoir le travail que nous venons de faire quatre fois, pour les valeurs réelles, imaginaires, positives et négatives.  $\square$

**Proposition 27.60** (Continuité de l'intégrale par rapport à la mesure[738]).

Soit  $p \in [1, \infty[$ . Nous considérons  $f \in L^p(\Omega, \mathcal{A}, \mu)$ . Alors pour tout  $\epsilon > 0$ , il existe un  $\delta > 0$  tel que

$$\int_A |f|^p d\mu < \epsilon \quad (27.282)$$

dès que  $\mu(A) < \delta$ .

*Démonstration.* Soit  $\epsilon > 0$ . La proposition 27.59 dit que les applications étagées sont denses dans  $L^p$ . Il existe donc une fonction étagée  $\phi \in L^p(\Omega, \mathcal{A}, \mu)$  telle que  $\|f - \phi\|_p < \epsilon$ . Étant donné que  $\phi$  est étagée, elle prend un nombre fini de valeurs et aucune de ces valeurs est  $\infty$ . Donc il existe  $M > 0$  tel que  $|\phi(x)| < M$  pour tout  $x \in \Omega$ .

Nous posons  $\delta = (\epsilon/2M)^p$ . Soit  $A \in \mathcal{A}$  tel que  $\mu(A) < \delta$ . Nous utilisons l'inégalité de Minkowski de la proposition 27.45(1) :

$$\|f\|_{A,p} \leq \|f - \phi\|_{A,p} + \|\phi\|_{A,p} \leq \frac{\epsilon}{2} + M\mu(A)^{1/p} < \epsilon. \quad (27.283)$$

Donc nous avons bien  $\int_A |f|^p < \epsilon$ .  $\square$

**Corolaire 27.61** ([1]).

Si  $f \in L^1([a, b])$ , alors

$$\lim_{h \rightarrow 0} \int_a^{a+h} f = 0. \quad (27.284)$$

*Démonstration.* Nous notons  $\lambda$  la mesure de Lebesgue. Soit  $\epsilon > 0$ . La proposition 27.60 dit qu'il existe  $\delta > 0$  tel que  $\lambda(A) < \delta$  implique  $\int_A f = 0$ . Si  $|h| < \delta$  nous avons  $\lambda([a, a+h]) < \delta$  et donc

$$\int_a^{a+h} f d\lambda = \int_{[a, a+h]} f d\lambda < \epsilon. \quad (27.285)$$

$\square$

**Proposition 27.62** ([1]).

Soit un intervalle  $I \subset \mathbb{R}$ . Soient  $f \in L^1(I)$  et  $a \in I$ . Alors l'application

$$\begin{aligned} \phi: I &\rightarrow \mathbb{R} \\ x &\mapsto \int_a^x f(t) dt \end{aligned} \quad (27.286)$$

est continue.

*Démonstration.* Soit  $h > 0$  tel que  $B(a, h) \subset I$ , étant compris que la boule est au sens de la topologie sur  $I$ ; donc si  $a$  est une extrémité de  $I$ , cela ne comprend que des  $h \geq 0$  ou  $h \leq 0$ . Nous avons

$$\phi(x+h) - \phi(x) = \int_a^{x+h} f - \int_a^x f = \int_x^{x+h} f. \quad (27.287)$$

La limite  $h \rightarrow 0$  donne zéro par le corolaire 27.61.  $\square$

### 27.3.10 Densité des fonctions infiniment dérivables à support compact

#### Définition 27.63.

Une fonction est **étagée par rapport à  $L^p$**  si elle est de la forme

$$f = \sum_{k=1}^N c_k \mathbb{1}_{B_k} \quad (27.288)$$

où les  $B_k$  sont des mesurables disjoints et  $\mathbb{1}_{B_k} \in L^p$  pour tout  $k$ .

LemWHIRdaX

#### Lemme 27.64.

Si  $f$  est une fonction étagée en même temps qu'être dans  $L^p$ , alors elle est étagée par rapport à  $L^p$ .

*Démonstration.* Nous pouvons écrire

$$f = \sum_{k=1}^N c_k \mathbb{1}_{B_k} \quad (27.289)$$

où les  $B_k$  sont disjoints. Par hypothèse  $\|f\|_p$  existe. Donc chacune des intégrales  $\int_{\Omega} |\mathbb{1}_{B_k}|^p$  doit exister parce que les  $B_k$  étant disjoints, nous pouvons inverser la norme et la somme ainsi que la somme et l'intégrale :

$$\int_{\Omega} |f|^p = \int_{\Omega} \sum_{k=1}^N |c_k \mathbb{1}_{B_k}(x)|^p dx = \sum_{k=1}^N \int_{\Omega} |c_k \mathbb{1}_{B_k}(x)|^p dx = \sum_{k=1}^N |c_k|^p \int_{\Omega} |\mathbb{1}_{B_k}(x)|^p dx. \quad (27.290)$$

□

L'ensemble  $C_c^\infty(\mathbb{R}^d)$  des fonctions de classe  $C^\infty$  et à support compact sur  $\mathbb{R}^d$  est souvent également noté  $\mathcal{D}(\mathbb{R}^d)$ .

ThoILGYXhX

#### Théorème 27.65 ([724]).

Nous avons des densités emboîtées. Ici  $D$  est un borélien borné de  $\mathbb{R}^d$  contenu dans  $B(0, r)$  et  $K$  est un compact contenant  $B(0, r + 2)$ .

(1) Les fonctions étagées par rapport à  $L^p$  sur  $\mathbb{R}^d$  sont denses dans  $L^p(\mathbb{R}^d)$ . A fortiori les fonctions étagées sont denses dans  $L^p$ , mais nous n'en aurons pas besoin ici.

ItemYVFFVrOlii

(2) Il existe une suite  $f_n$  dans  $C(K, \mathbb{C})$  telle que

$$f_n \xrightarrow{L^p} \mathbb{1}_D. \quad (27.291)$$

ItemYVFFVrOliii

(3) Si  $A$  est un borélien tel que  $\mathbb{1}_A \in L^p(\mathbb{R}^d)$ <sup>30</sup>, alors il existe une suite de boréliens bornée  $(D_n)_{n \in \mathbb{N}}$  tels que

$$\mathbb{1}_{D_n} \xrightarrow{L^p} \mathbb{1}_A. \quad (27.292)$$

ItemYVFFVrOliv

(4) Il existe une suite  $\varphi_n$  dans  $\mathcal{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$  telle que

$$\varphi_n \xrightarrow{L^p} \mathbb{1}_D. \quad (27.293)$$

ItemYVFFVrOIv

(5) L'ensemble  $\mathcal{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$  est dense dans  $L^p(\mathbb{R}^d)$  pour tout  $1 \leq p < \infty$ .

*Démonstration.* Nous allons montrer les choses point par point.

(1) Si  $f \in L^1(\mathbb{R}^d)$ , nous savons par le théorème 14.118 qu'il existe une suite  $f_n$  de fonctions étagées convergeant ponctuellement vers  $f$  telle que  $|f_n| \leq |f|$ . La proposition 27.25 nous dit qu'alors  $f_n \xrightarrow{L^p} f$ .

La fonction  $f_n$  étant étagée et dans  $L^p$  en même temps, elle est automatiquement étagée par rapport à  $L^p$  par le lemme 27.64.

<sup>30</sup>. Je pense que cette hypothèse manque dans [724]. En tout cas je vois mal comment je pourrais justifier les différentes étapes de la preuve en prenant par exemple  $A = \mathbb{R}^d$ .

ItemYVFVrOIi

- (2) C'est le théorème d'approximation 14.242 appliqué au borélien  $D$  contenu dans l'espace mesuré  $K$ .
- (3) En vertu du point (2), il existe  $f \in C^0(K, \mathbb{R})$  telle que

$$\|f - \mathbb{1}_D\|_p \leq \epsilon. \quad (27.294)$$

Ensuite, par le théorème de Weierstrass, il existe  $\varphi \in C^\infty(K, \mathbb{R})$  telle que  $\|f - \varphi\|_\infty \leq \epsilon$ . Nous avons aussi

$$\|\varphi - f\|_p^p = \int_K |\varphi(x) - f(x)|^p dx \leq \mu(K) \|\varphi - f\|_\infty^p \leq \epsilon^p \mu(K). \quad (27.295)$$

Quitte à prendre un  $\varphi$  correspondant à un  $\epsilon$  plus petit, nous avons

$$\|\varphi - f\| \leq \epsilon. \quad (27.296)$$

En combinant et en passant à  $\epsilon/2$  nous avons trouvé une fonction  $\varphi \in C^\infty(K, \mathbb{R})$  telle que

$$\|\varphi - \mathbb{1}_D\| \leq \epsilon. \quad (27.297)$$

- (4) Nous considérons les boréliens fermés  $D_n = A \cap B(0, n)$ . Alors  $\mathbb{1}_{D_n} \in L^p$  et nous avons pour  $n$  assez grand :

$$\int_{\mathbb{R}^d} |\mathbb{1}_{D_n}(x) - \mathbb{1}_A(x)|^p dx = \int_{\mathbb{R}^d \setminus B(0, n)} |\mathbb{1}_A(x)|^p dx < \epsilon, \quad (27.298)$$

c'est-à-dire que  $\mathbb{1}_{D_n} \xrightarrow{L^p} \mathbb{1}_A$ .

- (5) Il suffit de remettre tout ensemble. Si  $f \in L^p(\mathbb{R}^d)$ , par le point (2) nous commençons par prendre  $\sigma$  étagée par rapport à  $L^p$  telle que

$$\|\sigma - f\|_p \leq \epsilon. \quad (27.299)$$

Ensuite nous écrivons  $\sigma$  sous la forme

$$\sigma = \sum_{k=1}^N c_k \mathbb{1}_{B_k} \quad (27.300)$$

et nous appliquons le point (3) à chacune des  $\mathbb{1}_{B_k}$  pour trouver des boréliens bornés  $D_k$  tels que

$$\|\mathbb{1}_{D_k} - \mathbb{1}_{B_k}\|_p \leq \epsilon. \quad (27.301)$$

Enfin nous appliquons le point (4) pour trouver des fonctions  $\varphi_k \in C_c^\infty(\mathbb{R}^d)$  telles que

$$\|\varphi_k - \mathbb{1}_{D_k}\|_p \leq \epsilon. \quad (27.302)$$

Il n'est pas compliqué de calculer que

$$\left\| \sum_{k=1}^N c_k \varphi_k - f \right\|_p \leq 2\epsilon \sum_{k=1}^N c_k + \epsilon. \quad (27.303)$$

□

CorFZWooYNbtPz

### Corolaire 27.66.

Si  $1 < p < \infty$  alors la partie<sup>31</sup>  $L^2([0, 1]) \cap L^p([0, 1])$  est dense dans  $L^p([0, 1])$ .

31. Nous parlons bien ici de l'ensemble  $L^2$  parce que nous le considérons sans norme ou topologie particulière. La densité dont nous parlons ici est celle pour la topologie de  $L^p$ .

*Démonstration.* Nous savons du théorème 27.65(5) que  $C_c^\infty([0, 1])$  est dense dans  $L^p$ . Mais nous avons évidemment  $C_c^\infty \subset L^2 \cap L^p$ , donc  $L^2 \cap L^p$  est dense dans  $L^p$ .  $\square$

LemCUIJzkA

**Lemme 27.67** ([724, 691]).

Soit  $1 \leq p < \infty$  et  $f \in L^p(\Omega)$ . Nous notons  $\tau_v$  l'opérateur de translation par  $v$  :

$$\begin{aligned} \tau_v : L^p(\Omega) &\rightarrow L^p(\Omega) \\ f &\mapsto \left[ x \mapsto f(x - v) \right]. \end{aligned} \quad (27.304)$$

Pour chaque  $f \in L^p(\Omega)$ , l'application

$$\begin{aligned} \tau(f) : \mathbb{R}^d &\rightarrow L^p(\Omega) \\ v &\mapsto \tau_v(f) \end{aligned} \quad (27.305)$$

est continue en  $v = 0$ , c'est-à-dire

$$\lim_{v \rightarrow 0} \|\tau_v(f) - f\|_p = 0. \quad (27.306)$$

*Démonstration.* Nous commençons par supposer que  $f$  est dans  $\mathcal{D}(\Omega)$ , et nous verrons ensuite comment généraliser.

(1) Si  $f \in \mathcal{D}(\Omega)$

Soit une suite  $v_i \xrightarrow{\mathbb{R}^d} 0$ , et posons  $f_i = \tau_{v_i}(f)$ ; le but est de montrer que  $f_i \xrightarrow{L^p} f$ . Pour cela, la fonction  $f - f_i$  est également à support compact, et qui plus est, si  $\text{supp}(f) \subset B(0, r)$ , alors  $\text{supp}(f - f_i) \subset B(0, r + |v_i|)$ , et l'ensemble

$$S = \overline{B(0, r + \max_i |v_i|)} \quad (27.307)$$

est un compact contenant les supports de tous les  $f - f_i$ . Le maximum existe parce que  $v_i \rightarrow 0$ . Voilà qui « majore » le domaine de  $f - f_i$  uniformément en  $i$ .

Majorons maintenant  $|f - f_i|^p$  de façon uniforme en  $i$ . Soit le nombre

$$M = 2 \max_{x \in \mathbb{R}^d} \{f(x)\}. \quad (27.308)$$

La fonction qui vaut  $M^p$  sur  $S$  et zéro ailleurs est une fonction intégrable qui majore  $|f - f_i|^p$ . Nous pouvons donc utiliser la convergence dominée de Lebesgue (théorème 14.207) pour écrire

$$\lim_{i \rightarrow \infty} \|f - f_i\|_p^p = \lim_{i \rightarrow \infty} \int_{\Omega} |f(x) - f(x - v_i)|^p dx = \int_{\Omega} \lim_{i \rightarrow \infty} |f(x) - f(x - v_i)|^p dx = 0. \quad (27.309)$$

(2) Pour  $f \in L^p(\Omega)$

Soit  $\epsilon > 0$ ,  $f \in L^p(\Omega)$  et  $\varphi \in \mathcal{D}(\Omega)$  tel que  $\|f - \varphi\|_p \leq \epsilon$ . Cela est possible par la densité de  $\mathcal{D}(\Omega)$  dans  $L^p(\Omega)$  vue en 27.65(5). Nous choisissons de plus  $|v|$  assez petit pour avoir  $\|\tau_v(\varphi) - \varphi\|_p < \epsilon$ , qui est possible en vertu de ce que nous venons de démontrer à propos des fonctions à support compact. De plus  $\tau_v$  étant une isométrie de  $L^p$  nous avons  $\|\tau_v(\varphi) - \tau_v(f)\| = \|\varphi - f\| < \epsilon$ . Nous avons tout pour majorer :

$$\|f - \tau_v(f)\| \leq \|f - \varphi\| + \|\varphi - \tau_v(\varphi)\| + \|\tau_v(\varphi) - \tau_v(f)\| \leq 3\epsilon. \quad (27.310)$$

Nous avons donc bien  $\lim_{v \rightarrow 0} \|f - \tau_v(f)\| = 0$ .

 $\square$

### 27.3.11 Approximation

ThoJsBKir

**Théorème 27.68** ([368]).

Soit  $I$  un intervalle de  $\mathbb{R}$ . L'espace  $\mathcal{D}(I)$  des fonctions continues à support compact sur  $I$  est dense dans  $L^2(I)$ .

Ce théorème sera généralisé à tous les  $L^p(\mathbb{R}^d)$  par le théorème 27.65. Cependant  $L^p$  n'étant pas un Hilbert, il faudra travailler sans produit scalaire.

*Démonstration.* Soit  $g \in L^2(I)$  une fonction telle que  $g \perp f$  pour toute fonction  $f \in C_c(I)$ . Nous avons donc

$$\langle f, g \rangle = \int_I f \bar{g} = 0. \quad (27.311)$$

En passant éventuellement aux composantes réelles et imaginaires nous pouvons supposer que les fonctions sont toutes réelles. Nous décomposons  $g$  en parties positives et négatives :  $g = g^+ - g^-$ . Notre but est de montrer que  $g^+ = g^-$ , c'est-à-dire que  $g$  est nulle. La proposition 25.15 conclura que  $C_c(I)$  est dense dans  $L^2(I)$ .

Soit un intervalle  $[a, b] \subset I$  et une suite croissante de fonctions  $f_n \in C_c(I)$  qui converge vers  $\mathbb{1}_{[a,b]}$ . Par hypothèse pour chaque  $n$  nous avons

$$\int_I f_n g^+ = \int_I f_n g^-. \quad (27.312)$$

La suite étant croissante, le théorème de la convergence monotone (théorème 14.178) s'applique et nous avons

$$\lim_{n \rightarrow \infty} \int_I f_n g^+ = \int_a^b g^+, \quad (27.313)$$

de telle sorte que nous ayons, pour tout intervalle  $[a, b] \subset I$  l'égalité

$$\int_a^b g^+ = \int_a^b g^-. \quad (27.314) \quad \text{EqYIErAM}$$

De plus ces intégrales sont finies parce que

$$\int_a^b g^+ \leq \int_a^b |g| = \int_I |g| \mathbb{1}_{[a,b]} = \langle |g|, \mathbb{1}_{[a,b]} \rangle \leq \|g\|_{L^2} \sqrt{b-a} < \infty \quad (27.315)$$

par l'inégalité de Cauchy-Schwarz.

Soit maintenant un ensemble mesurable  $A \subset I$ . La fonction caractéristique  $\mathbb{1}_A$  est mesurable et il existe une suite croissante de fonctions étagées  $(\varphi_n)$  convergente vers  $\mathbb{1}_A$  par le théorème fondamental d'approximation 14.118. À multiples près, les fonctions  $\varphi_n$  sont des sommes de fonctions caractéristiques du type  $\mathbb{1}_{[a,b]}$ , par conséquent, en vertu de (27.314) nous avons

$$\int_I \varphi_n g^+ = \int_I \varphi_n g^-. \quad (27.316)$$

Une fois de plus nous pouvons utiliser le théorème de la convergence monotone et obtenir

$$\int_A g^+ = \int_A g^- \quad (27.317)$$

pour tout ensemble mesurable  $A \subset I$ . Si nous notons  $dx$  la mesure de Lebesgue, les mesures  $g^+ dx$  et  $g^- dx$  sont par conséquent égales et dominées par  $dx$ . Par le corolaire 14.236 du théorème de Radon Nikodym, les fonctions  $g^+$  et  $g^-$  sont égales.  $\square$



### 27.3.12 Partie uniformément intégrable

DefOZlZnse

#### Définition 27.69.

Une partie  $H \subset L^1(\Omega, \mu)$  est **uniformément intégrable** si

$$\lim_{a \rightarrow \infty} \left( \sup_{f \in H} \int_{|f| > a} |f(x)| d\mu(x) \right) = 0. \quad (27.318)$$

Notons dans cette définition que vu que  $f \in L^1$  nous avons toujours

$$\lim_{a \rightarrow \infty} \int_{|f| > a} |f(x)| d\mu(x) = 0. \quad (27.319)$$

LEMooUOQDooIkVyWm

#### Lemme 27.70 ([1]).

La partie  $H = \{f_n\}_{n \in \mathbb{N}}$  de  $L^1(\Omega, \mu)$  est uniformément intégrable<sup>32</sup> si et seulement si

$$\lim_{a \rightarrow \infty} \left( \limsup_n \int_{|f_n| > a} |f_n| \right) = 0. \quad (27.320)$$

Cette formule est prise comme définition de l'uniforme intégrabilité par [739].

*Démonstration.* Quand on a une suite  $(x_n)$  de réels positifs, nous avons toujours  $\limsup_n x_n \leq \sup_n x_n$ . □

#### ⚠ Avertissement/question au lecteur !! 27.71

Je n'ai pas vérifié l'énoncé de 27.72. Si vous pensez que c'est faux, écrivez-moi. Écrivez-moi également si vous avez une démonstration.

PROPooPZRAooSigKt

#### Proposition 27.72 ([1]).

Soit une partie uniformément intégrable  $\{f_n\}_{n \in \mathbb{N}}$ . Soit une fonction intégrable  $g$ .

- (1) La partie  $\{f_n + g\}$  est uniformément intégrable.
- (2) Pour tout  $\lambda \in \mathbb{R}$ , la partie  $\{\lambda f_n\}$  est uniformément intégrable.
- (3) Si pour tout  $n$  nous avons  $g_n \leq f_n$ , alors la partie  $\{g_n\}$  est uniformément intégrable.

PROPooPFZJooLySrgp

#### Proposition 27.73 ([1]).

Soit  $0 < r < \infty$ . Si  $f \in L^r(\Omega, \mathcal{A}, \mu)$ , alors

$$\int_{|f|^r \geq \lambda} |f|^r d\mu \xrightarrow{\lambda \rightarrow \infty} 0. \quad (27.321)$$

THOooGLYJooCbNysA

#### Théorème 27.74 ([740]).

Soient un espace mesuré fini  $(\Omega, \mathcal{A}, \mu)$  et une partie uniformément intégrable<sup>33</sup>  $H \subset L^1(\Omega)$ . Nous avons

ITEMooCAIFooHXOKQb

- (1)  $\sup_{f \in H} \|f\|_{L^1(\Omega)} < \infty$ .
- (2) Pour tout  $\epsilon > 0$ , il existe  $\delta_\epsilon > 0$  tel que

$$\int_A |f| d\mu < \epsilon \quad (27.322)$$

pour tout  $f \in H$  et pour tout  $A \in \mathcal{A}$  tel que  $\mu(A) < \delta_\epsilon$ .

32. Définition 27.69.

33. Définition 27.69.

## 27.4 Autres résultats de convergence

**Proposition 27.75** ([1]).

PROPooGJBIooAYsZQL

Soit  $0 < p < \infty$ . Si  $|f_n| \xrightarrow{L^p} 0$ , alors  $\liminf_n |f_n| = 0$  presque partout.

*Démonstration.* Nous y allons par contradiction en supposant qu'il existe une partie mesurable  $A$  de mesure non nulle telle que  $\liminf_n |f_n(\omega)| > 0$  pour tout  $\omega \in A$ .

Pour chaque  $\omega \in A$ , nous posons  $n_\omega \in \mathbb{N}$  et  $v_\omega > 0$  tels que pour tout  $n \geq n_\omega$ ,  $|f_n(\omega)| > v_\omega$ . Ensuite nous posons

$$A_{n,k} = \{\omega \in A \text{ tel que } n_\omega \leq n \text{ et } v_\omega > 1/k\}. \quad (27.323)$$

Nous avons l'union dénombrable  $A = \bigcup_{n,k \in \mathbb{N}^2} A_{n,k}$ . Vu qu'une union dénombrable d'ensembles de mesure nulle est de mesure nulle, il existe  $(n_0, k_0)$  tel que  $\mu(A_{n_0, k_0}) > 0$ . Pour tout  $\omega \in A_{n_0, k_0}$  et pour tout  $n > n_0$ , nous avons

$$|f_\omega(\omega)| > \frac{1}{k_0}. \quad (27.324)$$

Nous avons donc la minoration, pour tout  $n \geq n_0$ ,

$$\|f_n\|_p^p = \int_{\Omega} |f_n(\omega)|^p d\mu(\omega) \quad (27.325a)$$

$$\geq \int_{A_{n_0, k_0}} |f_n(\omega)|^p d\mu(\omega) \quad (27.325b)$$

$$\geq \int_{A_{n_0, k_0}} \frac{1}{k_0^p} d\mu(\omega) \quad (27.325c)$$

$$= \frac{1}{k_0^p} \mu(A_{n_0, k_0}) > 0. \quad (27.325d)$$

Nous ne pouvons donc pas avoir  $|f_n| \xrightarrow{L^p} 0$ . □

PROPooUMBQooAhUgHA

**Proposition 27.76** ([741]).

Si  $0 < p \leq 1$ , alors la formule

$$d_r(f, g) = \int_{\Omega} |f - g|^r d\mu \quad (27.326)$$

est une métrique invariante par translation.

Parfois il faut admettre que la théorie des probabilités a des notations pratiques. Par exemple si  $X_n \xrightarrow{L^p} X$  alors  $E(X_n) \rightarrow E(X)$ . C'est la proposition suivante.

PROPooXXOYooCUtDZf

**Proposition 27.77** ([742]).

Soient un espace mesuré  $(\Omega, \mathcal{A}, \mu)$  ainsi que  $0 < r < \infty$ . Nous supposons que  $f_n \xrightarrow{L^r} f$ . Alors

$$\lim_{n \rightarrow \infty} \int_{\Omega} |f_n|^r = \int_{\Omega} |f|^r. \quad (27.327)$$

*Démonstration.* Si  $r \geq 1$ , c'est facile. L'application  $f \mapsto \|f\|_r$  est une norme<sup>34</sup> et la proposition 7.213 s'applique. Donc

$$|\|f_n\|_r - \|f\|_r| \leq \|f - f_n\|_r. \quad (27.328)$$

Nous considérons  $0 < r \leq 1$ . Nous considérons la métrique  $d_r$  donnée par la proposition 27.76. Notre hypothèse de convergence signifie que  $d_r(f_n, f) \rightarrow 0$ . Nous utilisons l'inégalité triangulaire :

$$d_r(f_n, 0) \leq \underbrace{d_r(f_n, f)}_{\rightarrow 0} + d_r(f, 0) \rightarrow d(f, 0). \quad (27.329)$$

---

34. Proposition 27.19(2).

Nous avons donc<sup>35</sup>

$$\limsup_n d_r(f_n, 0) \leq d_r(f, 0). \quad (27.330)$$

D'autre part la proposition 27.75 nous dit que  $\liminf_n |f_n(\omega) - f(\omega)| = 0$  pour tout  $\omega \in \Omega \setminus Z$  avec  $\mu(Z) = 0$ . Nous avons donc aussi  $\liminf |f_n| = |f|$  (hors de  $Z$ ) et  $\liminf |f_n|^r = |f|^r$  (hors de  $Z$ ). Vu que  $Z$  est de mesure nulle, nous pouvons l'oublier quand nous faisons des intégrales sur  $\Omega$  :

$$d_r(f, 0) = \int_{\Omega} |f|^r \quad (27.331a)$$

$$= \int_{\Omega} \liminf_n |f_n|^r \quad (27.331b)$$

$$= \liminf \int_{\Omega} |f_n|^r \quad \text{lem Fatou 14.182} \quad (27.331c)$$

$$= \liminf d_r(f_n, 0) \quad (27.331d)$$

$$\leq \liminf \left( d_r(f_n, f) + d_r(f, 0) \right) \quad (27.331e)$$

$$= d_r(f, 0). \quad (27.331f)$$

Vu que les expressions tout à gauche et tout à droite sont égales, toutes les inégalités sont des égalités. En particulier

$$\liminf_n d_r(f_n, 0) = d_r(f, 0). \quad (27.332)$$

En combinant avec (27.330), et en nous souvenant qu'une limite sup est toujours plus grande qu'une limite inf,

$$\limsup d_r(f_n, 0) \leq d_r(f, 0) \leq \liminf d_r(f_n, 0) \leq \liminf d_r(f_n, 0). \quad (27.333)$$

Encore une fois toutes les inégalités sont des égalités et nous avons

$$\liminf d_r(f_n, 0) = \limsup d_r(f_n, 0) = d_r(f, 0). \quad (27.334)$$

Cela prouve que  $\lim_{n \rightarrow \infty} d_r(f_n, 0)$  existe et vaut  $d_r(f, 0)$  (lemme 9.46). Enfin nous pouvons faire le calcul

$$\int_{\Omega} |f|^r = d_r(f, 0) = \lim_{n \rightarrow \infty} d_r(f_n, 0) = \lim_{n \rightarrow \infty} \int_{\Omega} |f_n|^r. \quad (27.335a)$$

□

## 27.5 Convolution

### Définition 27.78.

Pour toutes fonctions  $f, g: \mathbb{R}^n \rightarrow \mathbb{C}$  et pour tout  $x \in \mathbb{C}$  tels que l'intégrale de droite ait un sens<sup>36</sup>, nous définissons

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x - y)dy. \quad (27.336)$$

L'éventuelle fonction  $f * g$  ainsi définie est le **produit de convolution** de  $f$  et  $g$ .

Le théorème qui permet de dire que le produit de convolution n'est pas tout à fait ridicule est le suivant.

### Théorème 27.79 ([523, 743]).

Soient  $f, g \in L^1(\mathbb{R}^d)$ .

35. Notons que  $x_n \leq y_n \rightarrow \lambda$  n'implique pas que  $x_n$  converge. Par contre ça implique que  $\limsup_n(x_n) \leq \lambda$ .

36. Attention divulgâchis : ce sera le cas pour  $f, g \in L^1(\mathbb{R}^n)$  par le théorème 27.79.

(1) Pour presque tout  $x \in \mathbb{R}^d$ , la fonction

$$\begin{aligned} h_x: \mathbb{R}^d &\rightarrow \mathbb{C} \\ y &\mapsto g(x-y)f(y) \end{aligned} \quad (27.337)$$

est dans  $L^1(\mathbb{R}^d)$ .

(2)  $f * g \in L^1(\mathbb{R}^d)$ .

(3)  $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$ .

*Démonstration.* Nous considérons l'application

$$\begin{aligned} \phi: \mathbb{R}^d \times \mathbb{R}^d &\rightarrow \mathbb{R}^d \times \mathbb{R}^d \\ (x, y) &\mapsto (x-y, y). \end{aligned} \quad (27.338)$$

Cela est un  $C^1$ -difféomorphisme dont le jacobien vaut

$$J_\phi(x, y) = \det \begin{pmatrix} (\partial_1 \phi_1)(x, y) & (\partial_2 \phi_1)(x, y) \\ (\partial_1 \phi_2)(x, y) & (\partial_2 \phi_2)(x, y) \end{pmatrix} = \det \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = 1. \quad (27.339)$$

C'est une première bonne chose.

Ensuite nous considérons la fonction

$$\begin{aligned} \alpha: \mathbb{R}^d \times \mathbb{R}^d &\rightarrow \mathbb{C} \\ (x, y) &\mapsto f(y)g(x). \end{aligned} \quad (27.340)$$

Par hypothèse, pour chaque  $y$ , la fonction  $x \mapsto |\alpha(x, y)|$  est dans  $L^1(\mathbb{R}^d)$ . Bref, le calcul suivant a un sens :

$$\int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} |\alpha| dx \right] dy = \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} |f(y)g(x)| dx \right] dy \quad (27.341a)$$

$$= \int_{\mathbb{R}^d} \left[ |f(y)| \int_{\mathbb{R}^d} |g(x)| dx \right] dy \quad (27.341b)$$

$$= \|f\|_{L^1(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)} \quad (27.341c)$$

$$< \infty. \quad (27.341d)$$

Le corolaire 14.304 nous dit alors que  $\alpha \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$ . Et notons au passage que

$$\|\alpha\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)} = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\alpha| = \|f\|_{L^1(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)} \quad (27.342)$$

parce que le théorème de Fubini 14.305 permet de scinder l'intégrale et de retomber sur (27.341c).

Vu que  $\alpha \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$ , nous pouvons utiliser le changement de variable 14.282 avec l'application  $\phi$  ci-dessus. En notant  $\lambda$  la mesure de Lebesgue,

$$\infty > \int_{\mathbb{R}^d \times \mathbb{R}^d} |\alpha| d\lambda = \int_{\mathbb{R}^d \times \mathbb{R}^d} (|\alpha| \circ \phi) |J_\phi| d\lambda \quad (27.343a)$$

$$= \int_{\mathbb{R}^d \times \mathbb{R}^d} |\alpha|(x-y, y) d\lambda(x, y) \quad (27.343b)$$

$$= \int_{\mathbb{R}^d \times \mathbb{R}^d} |f(y)g(x-y)| d\lambda(x, y) \quad (27.343c)$$

$$= \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} |f(y)g(x-y)| dy \right] dx. \quad (27.343d)$$

Nous avons scindé l'intégrale avec le théorème de Fubini pour la dernière étape.

Passons à l'intégrabilité de  $f * g$ . Nous avons

$$|(f * g)(x)| \leq \int_{\mathbb{R}^d} |f(y)g(x-y)| dy. \quad (27.344) \quad \text{EQooXMFYooKPIGfu}$$

Or nous venons de voir que (27.344) était, en tant que fonction de  $x$ , intégrable sur  $\mathbb{R}^d$  et que

$$\int_{\mathbb{R}^d} |(f * g)(x)| dx \leq \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} |f(y)g(x-y)| dy \right] dx = \int_{\mathbb{R}^d \times \mathbb{R}^d} |\alpha| d\lambda < \infty. \quad (27.345)$$

Cela prouve que  $f * g \in L^1(\mathbb{R}^d)$ , mais en nous souvenant de (27.342), cela prouve aussi que

$$\int_{\mathbb{R}^d} |(f * g)(x)| dx \leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |\alpha| d\lambda = \|f\|_{L^1(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)}, \quad (27.346)$$

c'est-à-dire que  $\|f * g\|_{L^2(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)} \|g\|_{L^1(\mathbb{R}^d)}$ . □

### Lemme 27.80.

*Le produit de convolution est commutatif : pour tout  $f, g \in L^1(\mathbb{R}^d)$  nous avons  $f * g = g * f$ .*

*Démonstration.* Le théorème de Fubini (théorème 14.305) permet d'écrire

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y) dy = \int_{-\infty}^{\infty} dy_1 \dots \int_{-\infty}^{\infty} dy_n f(y)g(x-y). \quad (27.347)$$

En effectuant le changement de variable  $z_i = x_i - y_i$  dans chacune des intégrales nous obtenons

$$(f * g)(x) = \int_{\mathbb{R}^n} g(z)f(x-z) dz = (g * f)(x). \quad (27.348)$$

Attention : on pourrait croire qu'un signe apparaît du fait que  $z = x - y$  donne  $dz = -dy$ . Mais en réalité, l'intégrale  $\int_{-\infty}^{+\infty}$  devient par le même changement de variables  $\int_{+\infty}^{-\infty}$  qui redonne un nouveau signe au moment de remettre dans l'ordre. □

### Lemme 27.81 ([744, 1]).

*Le produit de convolution est associatif sur  $L^1(\mathbb{R}^d)$ .*

*Démonstration.* Soient  $f, g, h \in L^1(\mathbb{R}^d)$ . L'existence de  $(f * g) * h$  ne fait pas de doute grâce au théorème 27.79. Nous avons d'abord

$$((f * g) * h)(u) = \int_{\mathbb{R}^d} (f * g)(x)h(u-x) dx \quad (27.349a)$$

$$= \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} f(y)g(x-y) dy \right] h(u-x) dx \quad (27.349b)$$

$$= \int_{\mathbb{R}^d} \left[ \int_{\mathbb{R}^d} f(y)g(x-y)h(u-x) dy \right] dx. \quad (27.349c) \quad \text{SUBEQooJKRBHooDeScXY}$$

Nous permutons les intégrales en suivant la procédure 14.306. Pour cela nous commençons par poser

$$s: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{C} \\ (x, y) \mapsto f(y)g(x-y)h(u-x), \quad (27.350)$$

et nous vérifions que  $|s|$  peut être successivement intégrée par rapport à  $y$  puis  $x$ . D'abord l'intégrale par rapport à  $y$  est

$$\int_{\mathbb{R}^d} |f(y)||g(x-y)| dy, \quad (27.351)$$

qui existe et qui vaut  $(|f| * |g|)(x)$  parce que  $|f|$  et  $|g|$  sont dans  $L^1(\mathbb{R}^d)$ . D'après le théorème 27.79, la fonction  $|f| * |g|$  est encore dans  $L^1(\mathbb{R}^d)$ . En ce qui concerne l'intégrale du résultat par rapport à  $x$ , nous avons

$$\int_{\mathbb{R}^d} (|f| * |g|)(x)|h(u-x)| dx, \quad (27.352)$$

qui existe et qui vaut  $((|f| * |g|) * |h|)(u)$ . Le corolaire 14.304 nous assure donc que  $s \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$ .

Nous permutons donc les intégrales dans (27.349c) pour obtenir

$$((f * g) * h)(u) = \int_{\mathbb{R}^d} f(y) \left[ \int_{\mathbb{R}^d} g(x - y) h(u - x) dx \right] dy. \quad (27.353)$$

Attardons-nous un instant sur l'intégrale interne, et utilisons l'invariance par translation de l'intégrale (lemme 14.257). Nous effectuons la translation  $x \rightarrow x + y$  :

$$\int_{\mathbb{R}^d} g(x - y) h(u - x) dx = \int_{\mathbb{R}^d} g(x) h(u - x - y) dx = \int_{\mathbb{R}^d} g(x) h((u - y) - x) dx \quad (27.354a)$$

$$= (g * h)(u - y). \quad (27.354b)$$

Nous pouvons reprendre notre calcul en (27.353) :

$$((f * g) * h)(u) = \int_{\mathbb{R}^d} f(y) \left[ \int_{\mathbb{R}^d} g(x - y) h(u - x) dx \right] dy \quad (27.355a)$$

$$= \int_{\mathbb{R}^d} f(y) (g * h)(u - y) dy \quad (27.355b)$$

$$= (f * (g * h))(u). \quad (27.355c)$$

C'est ce que nous voulions. □

### Proposition 27.82.

Le couple  $(L^1(\mathbb{R}^d), *)$  est une algèbre de Banach<sup>37</sup>.

*Démonstration.* Point par point.

- (1) **Algèbre** La définition d'une algèbre est 1.449. Les différents points sont dans la linéarité de l'intégrale.
- (2) **Commutative** C'est la proposition 27.80.
- (3) **Associative** C'est le lemme 27.81.
- (4) **Normé** L'espace  $L^1$  a une norme ; c'est la norme  $\|\cdot\|_{L^1}$ .
- (5) **Complet** C'est le théorème de Riesz-Fischer 27.56.

□

La proposition suivante est une conséquence de l'inégalité de Minkowski sous forme intégrale de la proposition 27.50.

### Proposition 27.83.

Si  $1 \leq p \leq \infty$  et si  $f \in L^p(\mathbb{R}^d)$  et  $g \in L^1(\mathbb{R}^d)$  alors

- (1)  $f * g \in L^p$
- (2)  $\|f * g\|_p \leq \|f\|_p \|g\|_1$ .

### Proposition 27.84 ([745]).

Si  $f \in L^1(\mathbb{R})$  et si  $g$  est dérivable avec  $g' \in L^\infty$ , alors  $f * g$  est dérivable et  $(f * g)' = f * g'$ .

*Démonstration.* La fonction qu'il faut intégrer pour obtenir  $f * g$  est  $f(t)g(x - t)$ , dont la dérivée par rapport à  $x$  est  $f(t)g'(x - t)$ . La norme de cette dernière est majorée (uniformément en  $x$ ) par  $G(t) = |f(t)| \|g'\|_\infty$ . La fonction  $f$  étant dans  $L^1(\mathbb{R})$ , la fonction  $G$  est intégrable et le théorème de dérivation sous l'intégrale (théorème 17.19) nous dit que  $f * g$  est dérivable et

$$(f * g)'(x) = \frac{d}{dx} \int_{\mathbb{R}} f(t)g(x - t)dt = \int_{\mathbb{R}} f(t)g'(x - t)dt = (f * g')(x). \quad (27.356)$$

□

---

37. Algèbre de Banach, définition 7.315.

CORooBSPNooFwYQrc

**Corolaire 27.85.**

Si  $f \in L^1(\mathbb{R}^d)$  et si  $g$  est de classe  $C^\infty$ , alors  $f * g$  est de classe  $C^\infty$ .

*Démonstration.* Il s'agit d'itérer la proposition 27.84. □

**27.5.1 Fonction plateau, lemme d'Urysohn**

PROPooQWNNooHmHRbk

**Proposition 27.86** ([1]).

Soient deux fonction  $f$  et  $g$  sur  $\mathbb{R}^n$  telles que le produit de convolution  $f * g$  ait un sens. Si le support de  $f$  est dans le compact  $K$  et le support de  $g$  est dans  $B(0, r)$ , alors

$$\text{supp}(f * g) \subset K + \overline{B(0, r)} \quad (27.357)$$

PROPooBOZIOoAhKbPs

**Proposition 27.87** (Urysohn, Fonctions plateau[286, 746]).

Soit un ouvert  $\Omega \subset \mathbb{R}^d$  ainsi qu'un compact  $K$  dans  $\Omega$ . Il existe une fonction  $f \in C^\infty(\mathbb{R}^d)$  telle que

- (1)  $\text{supp}(f)$  est compact dans  $\Omega$ ,
- (2)  $0 \leq f \leq 1$
- (3)  $f = 1$  sur un voisinage de  $K$ .

*Démonstration.* Nous posons  $\delta = d(K, \Omega^c)/3$ .

- (1) **Les applications  $\phi$  et  $\phi_\epsilon$**  Nous nous souvenons de l'application  $\xi$  de la proposition 15.173. Elle est de classe  $C^\infty$  à support compact dans  $B(0, 1)$ . Nous considérons<sup>38</sup>

$$\phi(x) = \frac{\xi(x)}{\int_{\mathbb{R}^n} \xi(x) dx}, \quad (27.358)$$

de telle sorte que  $\int_{\mathbb{R}^n} \phi = 1$ . Pour chaque  $\epsilon > 0$  nous posons également

$$\begin{aligned} \phi_\epsilon: \mathbb{R}^n &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{\epsilon^n} \phi(x/\epsilon). \end{aligned} \quad (27.359)$$

Par le lemme 15.181, nous avons

$$\int_{B(0, \epsilon)} \phi_\epsilon = \int_{\mathbb{R}^n} \phi_\epsilon = 1. \quad (27.360)$$

EQooDUOCooLKPmOZ

- (2) **La partie  $K_1$**  Nous posons

$$K_1 = \{x \in \mathbb{R}^n \text{ tel que } \phi(x, K) < \delta\}. \quad (27.361)$$

EQooLXRMooKxKHSE

La partie  $K_1$  est bornée parce que si  $K \subset B(0, r)$ , alors  $K_1 \subset B(0, r + \delta)$ .

La partie  $K_1$  est fermée. Pour ce voir, nous prouvons que  $K_1^c$  est ouvert. Soit  $y \in K_1^c$ , et prouvons qu'il existe une boule centrée en  $y$  et contenue dans  $K_1^c$ . Soit  $k \in K_1$  réalisant  $d(z, K_1) = d(z, k)$  ( $d$  est une fonction continue sur le compact  $K_1$ , elle atteint ses bornes). Soit  $z \in B(y, \alpha)$  où  $\alpha$  va être précisé de telle sorte à avoir  $B(y, \alpha) \subset K_1^c$ . Nous avons

$$\begin{aligned} d(z, k) &\geq |d(z, y) - d(y, k)| && \text{cf. justif.} && (27.362a) \\ &= d(y, k) - d(z, y) && \text{cf. justi.} && (27.362b) \\ &> \delta + s - \alpha && \text{cf. justif.} && (27.362c) \end{aligned} \quad (27.362d)$$

SUBEQooAKWEooGpnxaX

SUBEQooGTKYooEPMiZR

SUBSETooGVBFooIUCcKE

Justifications.

38. L'intégrale de  $\xi$  ne pose pas de problèmes parce qu'elle est continue sur le compact de son support, de telle sorte qu'elle y soit bornée.

- Pour (27.362a). Lemme 7.141.
- Pour (27.362b). Nous choisissons  $\alpha < d(y, k)$ .
- Pour 27.362c. Vu que  $d(y, k) > \delta$ , nous posons  $s > 0$  tel que  $d(y, k) = \delta + s$ .

Nous avons prouvé qu'il existe  $s > 0$  tel que pour tout  $\alpha > 0$  suffisamment petit, nous avons  $d(z, k) > \delta + s - \alpha$ . Nous en déduisons que  $d(z, k) > \delta$  et donc que  $z \in K_1^c$ .

Nous avons prouvé que  $K - 1$  est borné et fermé. Il est donc compact (théorème de Borel-Lebesgue 9.24).

- (3) **La fonction  $f$**  Nous posons  $f = \phi_\epsilon * \mathbb{1}_{K_1}$  où  $\mathbb{1}_{K_1}$  est la fonction indicatrice de  $K_1$  et  $*$  dénote le produit de convolution<sup>39</sup>. Vu que  $\phi_\epsilon$  est de classe  $C^\infty$ , la fonction  $f$  est également de classe  $C^\infty$  par le corolaire 27.85.
- (4) **À propos de support** La proposition 27.86 dit que le support de  $f$  est

$$\text{supp}(f) \subset K_1 + \overline{B(0, \epsilon)}. \quad (27.363)$$

Vue la définition 27.361, nous avons  $d(K_1, \Omega^c) \geq \delta$ , de telle sorte que, en prenant  $\epsilon < \delta$  nous avons

$$\text{supp}(f) \subset \Omega. \quad (27.364)$$

Le support de  $f$  est également borné parce qu'il est contenu dans le borné  $K_1 + \overline{B(0, \epsilon)}$ . Nous en déduisons que  $f$  est une application  $C^\infty$  à support compact dans  $\Omega$ .

- (5)  **$f(x)$  si  $x \in K$**  Soit  $x \in K$ . Nous avons, par définition,

$$f(x) = \int_{\mathbb{R}^n} \phi_\epsilon(y) \mathbb{1}_{K_1}(x - y) dy \quad (27.365a)$$

$$= \int_{\overline{B(0, \epsilon)}} \phi_\epsilon(y) \mathbb{1}_{K_1}(x - y) dy \quad \text{cf. justif. } \text{SUBEQooOKGFooDhOvrB} \quad (27.365b)$$

$$= \int_{\overline{B(0, \epsilon)}} \phi_\epsilon(y) dy \quad \text{cv. justif. } \text{SUEQooBRHEooAvoQdR} \quad (27.365c)$$

$$= 1 \quad \text{par (27.360)}. \quad (27.365d)$$

Justifications.

- Pour 27.365b. Le support de  $\phi_\epsilon$  est dans  $\overline{B(0, \epsilon)}$ .
- Pour 27.365c. Vu que  $x \in K$  et que  $|y| < \epsilon < \delta$ , nous avons

$$d(x - y, K) \leq d(x - y, x) \leq \delta, \quad (27.366)$$

et donc  $x - y \in K_1$ .

- (6)  **$0 \leq f \leq 1$**  Nous avons

$$|f(x)| \leq \int_{\mathbb{R}^n} |\phi_\epsilon(y) \mathbb{1}_{K_1}(x - y)| dy \quad (27.367a)$$

$$\leq \int_{\mathbb{R}^n} |\phi_\epsilon(y)| dy \quad (27.367b)$$

$$= 1. \quad (27.367c)$$

Note : nous avons majoré  $\mathbb{1}_{K_1}(x - y)$  par 1.

□

---

39. Définition 27.78.



### 27.5.2 Partition de l'unité

THOooQFCQooSlgLpz

**Théorème 27.88** (Partition de l'unité[286]).

Soient un compact  $K \subset \mathbb{R}^d$  et des ouverts  $\{\Omega_i\}_{i=1,\dots,n}$  recouvrant  $K$ . Alors il existe des fonctions  $\phi_i \in C^\infty(\mathbb{R}^d)$  ( $i = 1, \dots, n$ ) telles que

- (1)  $0 \leq \phi_k \leq 1$
- (2)  $\text{supp}(\phi_k) \subset \Omega_k$ ,
- (3)  $\sum_{k=1}^n \phi_k = 1$  sur un voisinage de  $K$ .

Ces fonctions  $\phi_i$  sont une **partition de l'unité** subordonnée aux ouverts  $\Omega_i$ .

*Démonstration.* Nous considérons des compacts  $\{K_i\}_{i=1,\dots,n}$  comme dans le lemme 7.359. Pour chaque  $i$  nous avons  $K_i \subset \Omega_i$ , de telle sorte que nous pouvions utiliser la proposition 27.87. Nous avons donc des fonctions  $\psi_j \in C^\infty(\mathbb{R}^d)$  telles que  $\text{supp}(\psi_j)$  est compact dans  $\Omega_j$ ,  $\psi_j \geq 0$  et  $\psi_j = 1$  sur  $V_j$  qui est un voisinage ouvert de  $K_j$ .

Nous posons  $V = V_1 \cup \dots \cup V_n$ . C'est un ouvert qui contient  $K$  parce que

$$K \subset \bigcup_j K_j \subset \bigcup_j \Omega_j \subset \bigcup_j V_j. \quad (27.368)$$

Nous faisons de même pour  $K$  lui-même. Il existe une fonction  $\theta \in C^\infty(\mathbb{R}^d)$  telle que

- $\text{supp}(\theta)$  est compact dans  $V$
- $\theta = 1$  sur un voisinage de  $K$ ,
- $0 \leq \theta \leq 1$ .

Vu que  $\psi_j = 1$  sur  $V_j$ , nous avons  $\sum_{j=1}^n \psi_j > 0$  sur  $V$ .

Nous prouvons à présent que

$$1 - \theta + \sum_{k=1}^n \psi_k > 0 \quad (27.369)$$

EQooGABMooZZmlyQ

sur  $\mathbb{R}^d$ .

- (1) **Si**  $x \in V$  Alors  $1 - \theta(x) \geq 0$  et donc

$$1 - \theta(x) + \sum_{k=1}^n \psi_k(x) \geq \sum_{k=1}^n \psi_k(x) > 0. \quad (27.370)$$

- (2) **Si**  $x \notin V$  Vu que le support de  $\theta$  est dans  $V$ , nous avons  $\theta(x) = 0$ . Quant aux fonctions  $\psi_k$ , elles font un peu ce qu'elles veulent, mais elles sont positives<sup>40</sup> et donc

$$1 - \theta(x) + \sum_{k=1}^n \psi_k(x) \geq 1 - \theta(x) = 1. \quad (27.371)$$

Vu que (27.369) est prouvé, nous pouvons poser

$$\phi_j = \frac{\psi_j}{1 - \theta + \sum_{k=1}^n \psi_k} \quad (27.372)$$

sans peur pour le dénominateur. Nous avons  $\phi_j \in C^\infty(\mathbb{R}^d)$ . Nous savons que  $\theta = 1$  sur un voisinage de  $K$ . Sur ce voisinage nous avons

$$\sum_{j=1}^n \phi_j(x) = \frac{\sum_j \psi_j(x)}{1 - \theta(x) + \sum_k \psi_k(x)} = 1. \quad (27.373)$$

De plus  $\phi_j \geq 0$  parce que chacun des  $\psi_j$  l'est et parce que nous avons montré que le dénominateur était toujours strictement positif.

En enfin,

$$\text{supp}(\phi_j) = \text{supp}(\psi_j) \subset \Omega_j. \quad (27.374)$$

Donc les fonctions  $\phi_j$  sont celles que dont nous avons besoin.  $\square$

40. Dans le Frido, « positif » signifie dans  $[0, \infty]$ .

### 27.5.3 Approximation de l'unité

LemDQEKNNf

#### Lemme 27.89.

Soit  $f \in L^2(I)$  telle que

$$\int_I f \varphi = 0 \quad (27.375)$$

pour toute fonction  $\varphi \in C_c^\infty(I)$ . Alors  $f = 0$  presque partout sur  $I$ .

*Démonstration.* Nous considérons la forme linéaire

$$\begin{aligned} \phi: L^2(I) &\rightarrow \mathbb{C} \\ g &\mapsto \langle f, g \rangle = \int_I f \bar{g}. \end{aligned} \quad (27.376)$$

Par densité<sup>41</sup> nous pouvons aussi considérer une suite  $(\varphi_n)$  dans  $C_c^\infty(I)$  convergeant dans  $L^2$  vers  $f$ . Alors nous avons pour tout  $n$  :

$$\langle f, \varphi_n \rangle = 0. \quad (27.377)$$

En passant à la limite,  $\langle f, f \rangle = 0$ , ce qui implique  $f = 0$  dans  $L^2$  et donc  $f = 0$  presque partout en tant que bonne fonction.  $\square$

Ce résultat est encore valable dans les espaces  $L^p$  (proposition 27.190), mais il demande le théorème de représentation de Riesz<sup>42</sup>.

DEFooEFGNooOREmBb

#### Définition 27.90 ([1, 724]).

Nous considérons  $\Omega = \mathbb{R}^d$  ou  $(S^1)^d$ . Une **approximation de l'unité** sur  $\Omega$  autour de  $a \in \Omega$  est une suite  $(\varphi_n)$  de fonctions à valeurs réelles dans  $L^1(\Omega)$  telle que

- (1)  $\sup_k \|\varphi_k\|_1 < \infty$ ,
- (2) pour chaque  $n$  nous avons  $\int_\Omega \varphi_n = 1$ ,
- (3) si  $V$  est un voisinage de  $a$ , alors

ITEMooGVRQooHDbrcf

$$\lim_{k \rightarrow \infty} \int_{\Omega \setminus V} |\varphi_k| = 0. \quad (27.378)$$

En pratique, nous allons, sur  $\mathbb{R}^d$  toujours considérer des approximations de l'unité autour de 0, même si nous ne le précisons pas. Vous noterez que dans le cas de  $S^1$ , le choix du « point de base » est plus arbitraire.

Ce sont des fonctions dont la masse vient s'accumuler autour de zéro. En effet quel que soit le voisinage  $B(0, \alpha)$ , si  $k$  est assez grand, il n'y a presque plus rien en dehors.

Pour le point ((3)), si  $\Omega$  est  $S^1$ , la mesure que nous considérons est  $\frac{dx}{2\pi}$ .

#### Exemple 27.91.

Une façon de construire une approximation de l'unité sur  $\mathbb{R}$  est de considérer une fonction  $\varphi \in L^1(\Omega)$  telle que  $\int \varphi = 1$  puis de poser

$$\varphi_k(x) = k^d \varphi(kx). \quad (27.379)$$

Ici,  $\Omega$  peut être  $\mathbb{R}$  ou  $S^1$ .  $\triangle$

Le lemme suivant permet de construire des approximations de l'unité intéressantes. Nous aurons une version pour  $S^1$  dans le lemme 27.126.

41. Théorème 27.65(5).

42. Théorème 27.187.

LemCNjIYhv

**Lemme 27.92** ([724]).

Soit une fonction  $\varphi$  continue et positive à support compact sur  $\mathbb{R}^d$  telle que  $\varphi(x) > \varphi(0)$  pour tout  $x \neq 0$ . Si nous posons

$$\varphi_n(x) = \left( \int \varphi(y)^n \right)^{-1} \varphi(x)^n, \quad (27.380)$$

alors la suite  $(\varphi_n)$  est une approximation de l'unité.

Voici un théorème qui donne les propriétés à propos du produit de convolution avec une approximation de l'unité dans  $\mathbb{R}^d$ . Une version pour  $S^1$  sera le théorème 27.127.

ThoYQbqEez

**Théorème 27.93** ([724]).

Soit  $(\varphi_k)$  une approximation de l'unité sur  $\mathbb{R}^d$ .

(1) Si  $g$  est mesurable et bornée sur  $\mathbb{R}^d$  et si  $g$  est continue en  $x_0$  alors

$$(\varphi_k * g)(x_0) \rightarrow g(x_0). \quad (27.381)$$

(2) Si  $g \in L^p(\mathbb{R}^d)$  ( $1 \leq p < \infty$ ) alors

$$\varphi_k * g \xrightarrow{L^p} g. \quad (27.382)$$

(3) Si  $g$  est uniformément continue et bornée, alors

$$\varphi_k * g \xrightarrow{L^\infty} g \quad (27.383)$$

*Démonstration.* En plusieurs points.

(1) Nous notons  $d_k = (\varphi_k * g)(x_0) - g(x_0)$  et nous devons prouver que  $d_k \rightarrow 0$ . Vu que  $\varphi_k$  est d'intégrale 1 sur  $\mathbb{R}^d$  nous pouvons écrire

$$d_k = \int_{\mathbb{R}^d} \varphi_k(y) g(x_0 - y) dy - \int_{\mathbb{R}^d} g(x_0) \varphi_k(y) dy, \quad (27.384)$$

et donc

$$|d_k| = \left| \int_{\mathbb{R}^d} (g(x_0 - y) - g(x_0)) \varphi_k(y) dy \right| \leq \int_{\mathbb{R}^d} |g(x_0 - y) - g(x_0)| |\varphi_k(y)| dy. \quad (27.385)$$

Nous notons  $M = \sup_k \|\varphi_k\|_1$ , et nous considérons  $\alpha > 0$  tel que

$$|g(x_0 - y) - g(x_0)| \leq \epsilon \quad (27.386)$$

pour tout  $y \in B(0, \alpha)$ . Nous nous restreignons maintenant aux  $k$  suffisamment grands pour que  $\int_{\mathbb{R}^d \setminus B(0, \alpha)} |\varphi_k(y)| dy \leq \epsilon$ . Alors en découpant l'intégrale en  $B(0, \alpha)$  et son complémentaire dans  $\mathbb{R}^d$ ,

$$|d_k| \leq \epsilon M + \int_{\mathbb{R}^d \setminus B(0, \alpha)} 2\|g\|_\infty |\varphi_k(y)| dy \leq \epsilon M + 2\|g\|_\infty \epsilon \leq \epsilon C. \quad (27.387)$$

Donc oui, nous avons  $|d_k| \rightarrow 0$ , et donc le premier point du théorème.

(2) Cette fois  $g \in L^p(\mathbb{R}^d)$  et nous cherchons à montrer que  $\|d_k\|_p \rightarrow 0$ . Encore qu'ici  $d_k$  soit défini à partir d'un représentant dans la classe de  $g$  et que d'ailleurs, nous allons travailler avec ce représentant.

D'abord nous développons un peu ce  $d_k$  :

$$\|d_k\|_p = \left[ \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} (g(x - y) - g(x)) \varphi_k(y) dy \right|^p dx \right]^{1/p} \quad (27.388a)$$

$$\leq \left[ \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |g(x - y) - g(x)| \cdot |\varphi_k(y)| dy \right)^p dx \right]^{1/p}. \quad (27.388b)$$

À cette dernière expression nous appliquons l'inégalité de Minkowski (théorème 27.45) sous la forme (27.208) pour la mesure  $d\nu(y) = |\varphi_k(y)|dy$  et  $f(x, y) = g(x - y) - g(x)$  :

$$\|d_k\|_p \leq \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} |g(x - y) - g(x)|^p dx \right)^{1/p} |\varphi_k(y)| dy = \int_{\mathbb{R}^d} \|\tau_y g - g\|_p |\varphi_k(y)| dy. \quad (27.389)$$

Par le lemme 27.67 nous pouvons trouver  $\alpha > 0$  tel que  $\|\tau_y g - g\|_p \leq \epsilon$  pour tout  $y \in B(0, \alpha)$ . Avec cela nous découpons encore le domaine d'intégration :

$$\|d_k\|_p \leq \int_{B(0, \alpha)} \underbrace{\|\tau_y g - g\|_p}_{\leq \epsilon} |\varphi_k(y)| dy + \int_{\mathbb{C}B(0, \alpha)} \underbrace{\|\tau_y g - g\|_p}_{\leq 2\|g\|_p} |\varphi_k(y)| dy \leq \epsilon M + 2\epsilon \|g\|_p. \quad (27.390)$$

- (3) Nous posons  $d_k(x) = (\varphi_k * g)(x) - g(x)$  et nous voulons prouver que  $\|d_k\|_\infty \rightarrow 0$ , c'est-à-dire que  $d_k(x)$  converge vers zéro uniformément en  $x$ . Nous posons aussi

$$\tau_y(g) : x \mapsto g(x - y). \quad (27.391)$$

En récrivant le produit de convolution, une petite majoration donne

$$|d_k(x)| \leq \int_{\mathbb{R}^d} \|\tau_y(g) - g\|_\infty |\varphi_k(y)| dy. \quad (27.392)$$

L'uniforme continuité de  $g$  signifie que pour tout  $\epsilon$ , il existe un  $\alpha$  tel que pour tout  $y \in B(0, \alpha)$ ,

$$\|\tau_y(g) - g\|_\infty \leq \epsilon. \quad (27.393)$$

Encore une fois nous découpons le domaine d'intégration en  $B = B(0, \alpha)$  et son complémentaire :

$$\|d_k\|_\infty \leq \int_B \underbrace{\|\tau_y(g) - g\|_\infty}_{\leq \epsilon} |\varphi_k(y)| dy + \int_{\mathbb{C}B} \underbrace{\|\tau_y(g) - g\|_\infty}_{\leq 2\|g\|_\infty} |\varphi_k(y)| dy \quad (27.394a)$$

$$\leq \epsilon M + 2\|g\|_\infty \epsilon \quad (27.394b)$$

où la seconde ligne est justifiée par le choix d'un  $k$  assez grand pour que  $\int_{\mathbb{C}B} |\varphi_k(y)| dy \leq \epsilon$ .

Nous avons donc bien  $\|d_k\|_\infty \rightarrow 0$ .

□

### Exemple 27.94.

Une petite remarque en passant : aussi triste que cela en ait l'air, la convergence uniforme n'implique pas la convergence  $L^p(\Omega)$  si  $\Omega$  n'est pas borné. En effet si  $f \in L^p$ , la suite donnée par

$$f_n(x) = f(x) + \frac{1}{n} \quad (27.395)$$

converge uniformément vers  $f$ , mais

$$\|f_n - f\|_p = \int_{\Omega} \frac{1}{n} \quad (27.396)$$

n'existe même pas si le domaine  $\Omega$  n'est pas borné.

△

## 27.6 Intégrale sur des variétés

### 27.6.1 Formes différentielles

Nous allons donner une toute petite introduction aux formes différentielles sur des variétés compactes.

LemdwLGFG

**Lemme 27.95** ([747]).

Soit  $\omega$  une  $k$ -forme sur  $\mathbb{R}^n$  et  $f$ , une fonction  $C^\infty$  sur  $\mathbb{R}^n$ . Alors  $d(f^*\omega) = f^*d\omega$ .

*Démonstration.* Nous effectuons la preuve par récurrence sur le degré de la forme. Soit d'abord une 0-forme, c'est-à-dire une fonction  $g: \mathbb{R}^n \rightarrow \mathbb{R}$ . Nous avons

$$d(d^*g)X = d(g \circ f)X = (dg \circ df)X = dg(dfX) = (f^*dg)(X). \quad (27.397)$$

Supposons maintenant que le résultat soit exact pour toutes les  $p-1$ -formes et montrons qu'il reste valable pour les  $p$ -formes. Par linéarité de la différentielle nous pouvons nous contenter de considérer la forme différentielle

$$\omega = g dx^1 \wedge \dots \wedge dx^p \quad (27.398)$$

où  $g$  est une fonction  $C^\infty$ . Pour soulager les notations nous allons noter  $dx^I = dx^1 \wedge \dots \wedge dx^{p-1}$ . Nous avons

$$d(f^*\omega) = d(f^*(g dx^I \wedge dx^p)) \quad (27.399a)$$

$$= d(f^*(g dx^I) \wedge f^* dx^p) \quad (27.399b)$$

$$= d(f^*(g dx^I)) \wedge f^* dx^p + (-1)^{p-1} f^*(g dx^I) \wedge (f^* dx^p) \quad (27.399c) \quad \text{gnAnSt}$$

$$= f^*(d(g dx^I)) \wedge f^* dx^p \quad (27.399d) \quad \text{xZrfZ}$$

$$= f^*(d(g dx^I) \wedge dx^p) \quad (27.399e)$$

$$= f^* d\omega \quad (27.399f) \quad \text{loWUi}$$

Justifications : (27.399c) est la formule de Leibniz. (27.399d) est parce que le second terme est nul :  $d(f^* dx^p) = f^*(d^2 x^p) = 0$ . Nous avons utilisé l'hypothèse de récurrence et le fait que  $d^2 = 0$ . L'étape (27.399f) est une utilisation à l'envers de la règle de Leibniz en tenant compte que  $d^2 x^p = 0$ .  $\square$

Soit  $M$  une variété de dimension  $n$  et  $\omega$  une  $n$ -forme différentielle

$$\omega_p = f(p) dx_1 \wedge \dots \wedge dx_n. \quad (27.400)$$

Si  $(U, \varphi)$  est une carte ( $U \subset \mathbb{R}^n$  et  $\varphi: U \rightarrow M$ ) alors nous définissons

$$\int_{\varphi(U)} \omega = \int_U f(\varphi(x)) dx_1 \dots dx_n. \quad (27.401)$$

Lorsque nous voulons intégrer sur une partie plus grande qu'une carte nous utilisons une partition de l'unité du théorème 27.88.

### 27.6.2 Intégrale d'une fonction sur une sous-variété

L'exemple typique est l'intégrale sur une surface dans  $\mathbb{R}^3$ , ou de volumes.

Nous supposons à présent que  $M$  est une variété compacte de dimension 2 dans  $\mathbb{R}^3$ . La compacité fait que  $M$  possède un atlas contenant un nombre fini de cartes  $F_i: W_i \rightarrow F_i(W_i)$ .

Si  $A \subset M$  est tel que pour chaque  $i$ ,  $A \cap F_i(W_i) = F_i(V_i)$  pour un ensemble  $V_i$  mesurable dans  $\mathbb{R}^2$ , alors nous considérons

$$A_1 = A \cap F_1(W_1) = F_1(V_1). \quad (27.402)$$

Ensuite, nous construisons  $A_2$  en considérant  $F_2(W_2)$  et en lui retranchant  $A_1$  :

$$A_2 = (A \cap F_2(W_2)) \setminus F_1(V_1). \quad (27.403)$$

En continuant de la sorte, nous construisons la décomposition

$$A = A_1 \cup \dots \cup A_p \quad (27.404)$$

de  $A$  en ouverts disjoints, chacun de ouverts  $A_p$  étant compris dans une carte.

Il est possible de prouver que dans ce cas, la définition suivante a un sens et ne dépend pas du choix de l'atlas effectué.

**Définition 27.96.**

Si  $f: A \rightarrow \mathbb{R}$  est une fonction continue, alors l'**intégrale** est le nombre

$$\int_A f = \sum_{i=1}^p \int_{A_i} f d\sigma_{F_i}. \quad (27.405)$$

**27.6.3 Densité des polynômes trigonométriques**

DEFooGCZAooFecAHB

**Définition 27.97.**

Le **système trigonométrique** donné par  $\{e_n\}_{n \in \mathbb{Z}}$  est

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}. \quad (27.406)$$

Une bonne partie de la douleur qu'évoque mot « densité » consiste à montrer que ce système est total dans  $L^2(S^1) = L^2([0, 2\pi])$ , et donc en est une base hilbertienne.

**Définition 27.98.**

Un **polynôme trigonométrique** est une fonction de la forme

$$P(t) = \sum_{n=-N}^N c_n e_n(t). \quad (27.407)$$

DEFooZDUKooRnYhhF

**Définition 27.99** (Coefficients de Fourier).

Pour toute fonction pour laquelle ça a un sens (que ce soit des fonctions  $L^2$  ou non), nous posons

$$c_k(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt. \quad (27.408)$$

EohIPoPH

Ces nombres sont les **coefficients de Fourier** de  $f$ .

Ces trois définitions n'ont à priori aucun rapport entre elles, et rien en particulier ne devrait vous faire penser à une égalité du type

$$f(x) = \sum_{n=-\infty}^{\infty} c_n(f) e_n(x). \quad (27.409)$$

Nous avons toutefois quelques liens.

LemZVfZlms

**Lemme 27.100.**

Deux petits résultats simples mais utiles à propos des polynômes trigonométriques.

(1) Si  $f \in L^1(S^1)$ , alors nous avons la formule

$$f * e_n = c_n(f) e_n. \quad (27.410)$$

(2) Si  $P$  est un polynôme trigonométrique et si  $f \in L^1(S^1)$  alors  $f * P$  est encore un polynôme trigonométrique.

*Démonstration.* Le premier point est un simple calcul :

$$(f * e_n)(x) = \int_0^{2\pi} f(x-t) e_n(t) dt \quad (27.411a)$$

En ce qui concerne le second point, nous notons  $P = \sum_{k=-N}^N P_k e_k$ , et par linéarité de la convolution,

$$f * P = \sum_{k=-N}^N P_k f * e_k = \sum_{k=-N}^n P_k c_k(f) e_k, \quad (27.412)$$

qui est encore un polynôme trigonométrique.  $\square$

ExDMnVSWF

**Exemple 27.101.**

Sur  $S^1$  nous construisons alors l'approximation de l'unité basée sur la fonction  $1 + \cos(x)$  et le lemme 27.92. Cette fonction est évidemment un polynôme trigonométrique parce que

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}. \quad (27.413)$$

Ensuite les puissances le sont aussi à cause de la formule du binôme :

$$(1 + \cos(x))^n = \sum_{k=0}^n \binom{n}{k} \cos^k(x), \quad (27.414)$$

dans laquelle nous pouvons remettre  $\cos(x)$  comme un polynôme trigonométrique et développer à nouveau la puissance avec (encore) la formule du binôme. La chose importante est qu'il existe une approximation de l'unité  $(\varphi_n)$  formée de polynômes trigonométriques.

Ce qui fait la spécificité des polynômes trigonométriques est qu'ils sont à la fois stables par convolution (lemme 27.100) et qu'ils permettent de créer une approximation de l'unité sur  $[0, 2\pi]$ . Ce sont ces deux choses qui permettent de prouver l'important théorème suivant.  $\triangle$

ThoQGPSSJq

**Théorème 27.102.**

Les polynômes trigonométriques sont dense dans  $L^p(S^1)$  pour  $1 \leq p < \infty$ .

*Démonstration.*

$$\varphi_k * f \xrightarrow{L^p} f \quad (27.415)$$

par le théorème 27.93. Nous avons donc convergence  $L^p$  d'une suite de polynômes trigonométriques, ce qui prouve que l'espace de polynômes trigonométriques est dense dans  $L^p(S^1)$ .  $\square$

**Remarque 27.103.**

Deux remarques.

- Il n'est pas possible que les polynômes trigonométriques soient dense dans  $L^\infty$  parce qu'une limite uniforme de fonctions continues est continue (c'est le théorème 12.381). Donc les polynômes trigonométriques ne peuvent engendrer que des fonctions continues.
- Nous donnerons au théorème 28.8 une démonstration indépendante de la densité des polynômes trigonométriques dans  $L^p(S^1)$ .

## 27.7 Espaces $L^2$ , généralités

SECooEVZSooLtLhUm

L'espace  $L^2$  est l'espace  $L^p$  défini en 27.10 avec  $p = 2$ . Cependant il possède une propriété extraordinaire par rapport aux autres  $L^p$ , c'est que la norme  $|\cdot|_2$  dérive d'un produit scalaire. Il sera donc un espace de Hilbert.

NORMooUEIEooYtlFse

**27.104.**

Nous en rappelons la construction. Soit  $(\Omega, \mathcal{A}, \mu)$  un espace mesuré. Nous considérons l'opération

$$\langle f, g \rangle = \int_{\Omega} f(\omega) \overline{g(\omega)} d\mu(\omega) \quad (27.416)$$

DefProdScallubrgTj

et la norme associée

$$\|f\|_2 = \sqrt{\langle f, f \rangle}. \quad (27.417)$$

Nous considérons l'ensemble

$$\mathcal{L}^2(\Omega, \mu) = \{f: \Omega \rightarrow \mathbb{C} \text{ tel que } \|f\|_2 < \infty\} \quad (27.418)$$

et la relation d'équivalence  $f \sim g$  si et seulement si  $f(x) = g(x)$  pour  $\mu$ -presque tout  $x$ .

Et enfin, nous considérons le quotient

$$L^2(\Omega, \mu) = \mathcal{L}^2(\Omega, \mu) / \sim. \quad (27.419)$$

LemIVWooZyWodb

**Lemme 27.105.**

Soit un espace mesuré<sup>43</sup>  $(\Omega, \mathcal{A}, \mu)$ .

(1) Pour tout  $f, g \in L^2(\Omega, \mathcal{A}, \mu)$ , le produit

$$\langle f, g \rangle = \int_{\Omega} f \bar{g} d\mu \quad (27.420)$$

EQooGLVUooObPmaX

est bien défini et est un nombre complexe<sup>44</sup>.

(2) L'opération  $(f, g) \mapsto \langle f, g \rangle$  est un produit hermitien<sup>45</sup>.

(3) Le couple  $(L^2(\Omega, \mathcal{A}, \mu), \langle \cdot, \cdot \rangle)$  est un espace de Hilbert<sup>46</sup>.

*Démonstration.* Que  $L^2(\Omega)$  soit un espace vectoriel est un cas particulier de la proposition 27.19. Voyons cette histoire de produit scalaire.

(1) **Pour de vraies fonctions** Nous commençons par analyser l'intégrale (27.420) dans le cas où  $f$  et  $g$  sont des fonctions, c'est-à-dire des représentants d'éléments de  $L^2$ .

Dans ce cas, l'inégalité de Hölder (proposition 27.39) avec  $p = q = 2$  nous indique que le produit  $f\bar{g}$  est un élément de  $L^1$ . Par conséquent la formule a un sens.

(2) **Passage aux classes**

Ensuite nous montrons que la formule passe au quotient. Pour cela, nous considérons des fonctions  $\alpha$  et  $\beta$  nulles presque partout et nous regardons le produit de  $f_1 = f + \alpha$  par  $g_1 = g + \beta$  :

$$\langle f_1, g_1 \rangle = \int f \bar{g} + \bar{\beta} f + \alpha \bar{g} + \alpha \bar{\beta}. \quad (27.421)$$

Les fonctions  $\bar{\beta}f$ ,  $\alpha\bar{g}$  et  $\alpha\bar{\beta}$  étant nulles presque partout, leur intégrale est nulle et nous avons bien  $\langle f_1, g_1 \rangle = \langle f, g \rangle$ . Nous pouvons donc considérer le produit sur l'ensemble des classes.

(3) **Produit hermitien** Pour vérifier que la formule est un produit hermitien, le seul point non évident est de prouver que  $\langle f, f \rangle = 0$  implique  $f = 0$ . Cela découle du fait que

$$\langle f, f \rangle = \int_{\Omega} |f|^2. \quad (27.422)$$

La fonction  $x \mapsto |f(x)|^2$  vérifie les hypothèses du lemme 14.199. Par conséquent  $|f(x)|^2$  est presque partout nulle.

(4) **Espace de Hilbert** En ce qui concerne le fait que  $L^2(\Omega)$  soit un espace de Hilbert, il s'agit simplement de se remémorer que c'est un espace complet (théorème 27.55) et dont la norme dérive d'un produit scalaire ou hermitien. Nous sommes donc bien dans la définition 25.3.

□

**27.106.**

Ces espaces seront utilisés pour de nombreuses applications. Nous en aurons besoin pour plusieurs combinaisons d'ensembles  $\Omega$  et de mesures  $\mu$ .

- Pour  $\mathbb{R}^d$
- Pour  $S^1$
- Pour  $[a, b]$
- Pour  $[0, 2\pi[$

43. Est-ce qu'il ne faudrait pas un peu plus d'hypothèses, comme  $\sigma$ -fini par exemple? Vérifiez et écrivez-moi quand vous avez la réponse.

44. Par opposition au fait que ce serait l'infini.

45. Définition 8.172. Pour rappel, nous considérons des fonctions à valeurs complexes. Si au contraire nous avions considéré seulement des fonctions à valeurs réelles, nous aurions eu un produit scalaire.

46. Définition 25.3.



— Pour  $[-T, T[$

Le premier est non compact et il est raisonnable de penser qu'il sera fondamentalement différent des autres. À isomorphismes assez triviaux près, les espaces des fonctions sur les trois autres sont identiques. Nous nous attendons donc à ce qu'ils aient les mêmes propriétés. Notons que du point de vue de  $L^2$ , étant donné qu'il y a un quotient par les parties de mesures nulles, prendre  $]0, 2\pi[$  ou  $[0, 2\pi]$  ou n'importe quelle autre possibilité de ce genre revient au même.

Afin de pouvoir utiliser ces espaces de façon optimale, et entre autres y définir les séries de Fourier, nous avons besoin, pour chacun d'entre eux de définir les éléments suivants :

- mesure
- produit de convolution
- le système trigonométrique (que nous allons montrer être une base hilbertienne)
- coefficients de Fourier

Ça fait pas mal de choses à définir. Il n'est pas besoin de définir un produit scalaire parce que le lemme 27.105 nous en donne un générique.

Les définitions qui viennent sont à prendre « tant que les formules ont un sens ». Nous parlons donc de fonctions dans  $\text{Fun}(\Omega, \mathbb{C})$ , l'ensemble de toutes les fonctions sur  $\Omega$  à valeurs dans  $\mathbb{C}$ . Nous verrons plus tard les espaces de fonctions sur lesquels tout a un sens.

## 27.8 L'espace $L^2(\mathbb{R}^d)$

La mesure est celle de Lebesgue. Le produit de convolution est donné, pour  $f, g \in \text{Fun}(\mathbb{R}^d, \mathbb{C})$ , par

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x - y)dy \quad (27.423)$$

Certaines de ses propriétés ont déjà été vues dans le théorème 27.79.

En ce qui concerne le système trigonométrique, pour tout  $\xi \in \mathbb{R}^d$  nous définirions bien

$$e_\xi(x) = e^{i\xi \cdot x}, \quad (27.424)$$

genre pour faire que les transformations de Fourier sont des séries continues ...mais bon. Nous n'allons pas tenter le diable plus que ça, et nous ne définissons

- pas de système trigonométrique,
- pas de coefficients de Fourier non plus,
- pas de théorie des séries de Fourier sur  $\mathbb{R}^d$ .

Quand je disais que la non-compactité de  $\mathbb{R}^d$  allait un peu changer les choses par rapport aux autres, je ne rigolais pas.

## 27.9 L'espace $L^2(S^1)$

L'espace  $S^1$  sera fait avec forces détails, parce qu'il va servir de base pour les espaces  $L^2([0, 2\pi[)$ ,  $L^2([-T, T[)$  ainsi que pour l'étude des fonctions périodiques sur  $\mathbb{R}$ .

En tant qu'ensemble,

$$S^1 = \{e^{it}\}_{t \in \mathbb{R}}, \quad (27.425)$$

sans garanties que ce paramétrage soit une bijection.

Il y a essentiellement deux façons de définir une intégrale sur  $S^1$ .

- (1) Voir  $S^1$  comme une sous-variété de  $\mathbb{R}^2$  et utiliser la définition 20.9. Cette façon a cependant deux inconvénients :

- Elle ne donne pas la tribu des mesurables sur  $S^1$ , c'est-à-dire que cette méthode ne donne pas de façon évidente une théorie de la mesure sur  $S^1$ .

— Il faut au moins deux cartes pour paramétrer le cercle. La fainéantise nous prévient que ça va être technique.

- (2) Rapporter la structure d'espace mesuré de  $[0, 2\pi[$  vers  $S^1$ , de force via le premier difféomorphisme qui nous passe par la tête, à savoir  $t \mapsto e^{it}$ .

Nous allons choisir la seconde possibilité, en gardant en tête qu'elle fonctionne de façon très simple un peu par coup de chance, voir la remarque 14.285(5).

### 27.9.1 Espace mesuré

Plusieurs choses sont déjà faites.

- Les boréliens de  $S^1$  sont décrits dans la proposition 18.71,
- la tribu de Lebesgue de  $S^1$  est décrite dans la proposition 18.73. Non, ce n'est pas la tribu induite de la tribu de Lebesgue de  $\mathbb{C}$ .

**Proposition 27.107** (Espaces de fonctions sur  $S^1$ [1]).

Soit l'espace mesuré  $(S^1, \mathcal{L}eb(S^1), \mu)$ .

- (1) La formule

$$\langle f, g \rangle = \int_{S^1} f \bar{g} d\mu \quad (27.426)$$

est un produit hermitien<sup>47</sup> sur  $L^2(S^1, \mathcal{L}eb(S^1), \mu)$ .

- (2) L'espace  $L^2(S^1)$  est un espace de Hilbert.

- (3) L'application<sup>48</sup>

$$\begin{aligned} \phi: L^2(S^1) &\rightarrow L^2([0, 2\pi[) \\ f &\mapsto \frac{1}{\sqrt{2\pi}} f \circ \varphi. \end{aligned} \quad (27.427)$$

est une bijection isométrique (isomorphisme d'espaces de Hilbert)

Tout cela se résume par l'égalité

$$L^2(S^1, \mathcal{L}eb(S^1), \mu) = L^2([0, 2\pi[, \mathcal{L}eb(\mathbb{R}), \lambda) \quad (27.428)$$

où nous avons fait un minuscule abus de notations : ici  $\mathcal{L}eb(\mathbb{R})$  est en réalité la tribu induite sur  $]0, 2\pi[$ .

*Démonstration.* Le fait que la formule (27.426) donne bien un produit hermitien est le lemme 27.105. Ce même lemme assure que le tout donne un espace de Hilbert.

Il nous reste à prouver le point (3). En ce qui concerne l'isométrie, nous posons<sup>49</sup>

$$\begin{aligned} \phi: L^2(S^1) &\rightarrow L^2([0, 2\pi[) \\ f &\mapsto \frac{1}{\sqrt{2\pi}} f \circ \varphi \end{aligned} \quad (27.429)$$

- (1) **Injection** Si  $\phi(f) = \phi(g)$ , alors pour tout  $x \in [0, 2\pi[$  nous avons  $f(\varphi(x)) = g(\varphi(x))$ . Vu que  $\varphi: [0, 2\pi[ \rightarrow S^1$  est une bijection nous avons alors  $f(s) = g(s)$  pour tout  $s \in S^1$ .
- (2) **Surjection** Si  $f \in L^2([0, 2\pi[)$ , nous posons  $g: S^1 \rightarrow \mathbb{C}$  par

$$g(s) = \sqrt{2\pi} f(\varphi^{-1}(s)). \quad (27.430)$$

Nous avons alors bien  $\phi(g)(x) = f(x)$ .

47. Définition 8.172.

48. Ici  $\varphi: [0, 2\pi[ \rightarrow S^1$  est la bijection usuelle de la proposition 18.62(1).

49. Notez que cette définition passe aux classes. Nous le répéterons pas.

(3) **Isométrie** Nous montrons que  $\phi$  préserve le produit scalaire : SUBEQSooRYYHooPcLXHN

$$\langle \phi(f), \phi(g) \rangle = \int_0^{2\pi} \phi(f)(x) \overline{\phi(g)(x)} d\lambda(x) \quad (27.431a)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} (f \circ \varphi)(x) \overline{(g \circ \varphi)(x)} d\lambda(x) \quad (27.431b)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} (f\bar{g}) \circ \varphi d\lambda \quad (27.431c)$$

Pour la suite nous devons invoquer la proposition 14.284 qui nous permet de changer une intégrale sur  $([0, 2\pi[, \mathcal{L}eb([0, 2\pi[), \lambda)$  en une intégrale sur  $(S^1, \mathcal{L}eb(S^1), \mu)$ . La première condition de cette proposition est que  $\mathcal{L}eb(S^1) = \varphi(\mathcal{L}eb([0, 2\pi[))$ . Cela est la proposition 18.73(2). La condition sur la mesure dans la proposition 14.284 n'est vraie ici qu'à un facteur  $2\pi$  près. Nous avons :

$$\int_{[0, 2\pi[} f d\lambda = 2\pi \int_{S^1} (f \circ \varphi^{-1}) d\mu. \quad (27.432)$$

Nous continuons le calcul (27.431) :

$$\langle \phi(f), \phi(g) \rangle = \frac{1}{2\pi} \int_0^{2\pi} (f\bar{g}) \circ \varphi d\lambda = \int_{S^1} f\bar{g} d\mu = \langle f, g \rangle. \quad (27.433)$$

□

## 27.9.2 Topologie

Nous considérons sur  $S^1$  la topologie induite de  $\mathbb{C}$ . Vu que  $S^1$  est fermé et borné dans  $\mathbb{C}$ , il en est une partie compacte. Par le lemme 7.112, l'espace  $S^1$  muni de sa topologie est un espace topologique compact.

Nous pouvons donc sans crainte affirmer que toute fonction continue  $f: S^1 \rightarrow \mathbb{K}$  est bornée et atteint ses bornes.

PROFoooEQDBooDfOrTZ

### Proposition-Définition 27.108.

Soit la formule

$$d(e^{ix}, e^{iy}) = \inf_{k \in \mathbb{Z}} |x - y + 2k\pi|. \quad (27.434)$$

- (1) Elle est bien définie (ne dépend pas des choix de  $x$  et  $y$  donnant les mêmes points dans  $S^1$ )
- (2) L'infimum est en réalité un minimum : il est atteint par un certain  $k \in \mathbb{Z}$  (qui, lui, dépend des choix).
- (3) La formule définit une distance<sup>50</sup> sur  $S^1$ .

Nous considérons sur  $S^1$  la topologie  $\tau_d$  découlant de cette distance.

*Démonstration.* Point par point.

- (1) Soient  $x', y' \in \mathbb{R}$  tels que  $e^{ix'} = e^{ix}$  et  $e^{iy'} = e^{iy}$ . Alors  $x' = x + 2l\pi$  et  $y' = y + 2l'\pi$  pour certains entiers  $l, l' \in \mathbb{Z}$  (corolaire 18.25). Nous avons alors  $|x' - y' + 2k\pi| = |x - y + 2\pi(k + l - l')|$  et

$$\inf_{k \in \mathbb{Z}} |x' - y' + 2k\pi| = \inf_{k \in \mathbb{Z}} |x - y + 2k\pi|. \quad (27.435)$$

- (2) Quels que soient  $x$  et  $y$  fixés, nous avons

$$\lim_{k \rightarrow \pm\infty} |x - y + 2k\pi| = \infty. \quad (27.436)$$

Donc l'infimum est forcément atteint par un  $k \in \mathbb{Z}$ .

- (3) Pour la distance, il y a plusieurs points à prouver.

---

50. Définition 7.139.

- Pour tout  $z, z' \in S^1$  nous avons  $d(z, z') \geq 0$  parce que la distance est donnée par une valeur absolue.
- Si  $d(z, z') = 0$ , alors il existe  $k$  tel que  $x = y + 2k\pi$ . Alors  $e^{ix} = e^{i(y+2k\pi)} = e^{iy}e^{2ki\pi} = e^{iy}$ . C'est-à-dire  $z = z'$ .
- Pour la symétrie, nous avons

$$|x - y + 2k\pi| = |y - x - 2k\pi| = |y - x + 2k'\pi| \quad (27.437)$$

en posant  $k' = -k$ . L'infimum étant pris sur  $k \in \mathbb{Z}$ , nous avons la symétrie  $d(e^{ix}, e^{iy}) = d(e^{iy}, e^{ix})$ .

- Pour attaquer l'inégalité triangulaire, nous considérons  $z_1 = e^{ix_1}$ ,  $z_2 = e^{ix_2}$  et  $z_3 = e^{ix_3}$ . Nous posons également  $k_1, k_2, k_3 \in \mathbb{Z}$  tels que  $d(z_1, z_3) = |x_1 - x_3 + 2k_1\pi|$ ,  $d(z_1, z_2) = |x_1 - x_2 + 2k_2\pi|$  et  $d(z_2, z_3) = |x_2 - x_3 + 2k_3\pi|$ . Nous avons alors

$$d(z_1, z_3) = \inf_{k \in \mathbb{Z}} |x_1 - x_3 + 2k\pi| \quad (27.438a)$$

$$= \inf_{k \in \mathbb{Z}} |x_1 - x_2 + x_2 - x_3 + 2k\pi| \quad (27.438b)$$

$$= \inf_{k \in \mathbb{Z}} |(x_1 - x_2 + 2k_2\pi) + (x_2 - x_3 + 2k_3\pi) + 2k\pi| \quad (27.438c)$$

$$\leq |(x_1 - x_2 + 2k_2\pi) + (x_2 - x_3 + 2k_3\pi)| \quad (27.438d)$$

$$\leq |x_1 - x_2 + 2k_2\pi| + |x_2 - x_3 + 2k_3\pi|. \quad (27.438e)$$

Notez que pour (27.438d), en posant  $k = 0$  nous avons bien une inégalité. En y pensant,  $k = 0$  réalise effectivement l'infimum, mais ce n'est pas indispensable.

□

Le cercle est bien connu pour être symétrique et en particulier avoir une symétrie sous les rotations. Nous allons voir quelques résultats qui vont dans le sens de dire que la distance définie sur  $S^1$  respecte cette symétrie.

LEMooCQC AooAEctbe

### Lemme 27.109.

Plusieurs points à propos de l'invariance de la topologie sous les rotations.

- (1) La distance est invariante sous les rotations, c'est-à-dire que si  $a, b \in S^1$  et si  $s \in \mathbb{R}$ , alors

$$d(e^{is}a, e^{is}b) = d(a, b). \quad (27.439)$$

ITEMooCIPYooTyPQLj

- (2) Les boules sont préservées sous les rotations<sup>51</sup>, c'est-à-dire que

$$e^{is}B_d(a, r) = B_d(e^{is}a, r). \quad (27.440)$$

- (3) La topologie est invariante sous les rotations :  $e^{is}\tau_d = \tau_d$ .

*Démonstration.* Point par point.

- (1) Si  $a = e^{ix}$  et  $b = e^{iy}$ , nous avons

$$d(a, b) = \inf_{k \in \mathbb{Z}} |x - y + 2k\pi| = \inf_{k \in \mathbb{Z}} |(x - s) - (y - s) + 2k\pi| = d(e^{is}a, e^{is}b). \quad (27.441)$$

Et de un.

- (2) Il faut une inclusion dans chaque sens.

- (2a)  $e^{is}B_d(a, r) \subset B_d(e^{is}a, r)$  Soit  $b \in e^{is}B_d(a, r)$ . Alors  $b = e^{is}b'$  pour un certain  $b' \in B_d(a, r)$ . Nous avons alors, en utilisant le premier point,

$$d(b, ae^{is}) = d(e^{-is}b, a) = d(b', a) < r. \quad (27.442)$$

Donc  $b \in B_d(e^{is}a, r)$ .

---

51. Pas chaque boule séparément, mais l'ensemble des boules

- (2b)  $\overline{B_d(e^{is}a, r)} \subset e^{is}B_d(a, r)$  Soit  $b \in B_d(e^{is}a, r)$ . Nous devons prouver que  $b \in e^{is}B_d(a, r)$ , c'est-à-dire que  $b = e^{is}b'$  pour un certain  $b' \in B_d(a, r)$  ou encore que  $e^{-is}b \in B_d(a, r)$ . En utilisant encore le premier point,

$$d(e^{-is}b, a) = d(b, e^{is}a) < r. \quad (27.443)$$

Donc oui,  $e^{-is}b \in B_d(a, r)$ .

- (3) Soit  $A \in \tau_d$ . Si  $a \in e^{is}A$ , alors  $a = e^{is}a'$  pour un certain  $a' \in A$ . Notre but est de prouver que  $e^{is}A$  contient un voisinage de  $a$ .

Vu que  $a' \in A$ , il existe  $r > 0$  tel que  $B_d(a', r) \subset A$ . Nous avons alors

$$e^{is}aB_d(a', r) \subset e^{is}A, \quad (27.444)$$

et comme  $e^{is}B_d(a', r) = B_d(e^{is}a', r) = B_d(a, r)$  nous avons bien

$$B_d(a, r) \subset e^{is}A. \quad (27.445)$$

□

Nous allons voir maintenant quelques résultats à propos de  $B_d(1, r)$  qui a la bonne figure d'être un ouvert qui s'étale symétriquement en partant de 1 (le point le plus à droite du cercle). Par rapport à la figure 27.1, il s'agit ni plus ni moins que de voir qu'une boule de rayon  $r$  autour de 1 est bien la partie indiquée (symétrique par rapport à 1 et de longueur d'arc  $r$  des deux côtés). De plus, ce voisinage n'est autre que la partie du cercle située à droite de la ligne en pointillés.

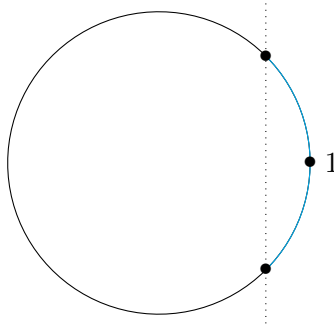


FIGURE 27.1: Un voisinage de 1 dans  $S^1$ .

LabelFigJQqVoolPTsYPZK

Ce lemme-ci montre que  $B_d(1, r)$  est une partie de  $S^1$  qui s'étale symétriquement autour de 1.

LEMooMYNVoolIWWsiV

### Lemme 27.110.

Soit l'application

$$\begin{aligned} \varphi: \mathbb{R} &\rightarrow S^1 \\ x &\mapsto e^{ix}. \end{aligned} \quad (27.446)$$

Nous avons  $B_d(1, r) = \varphi(]-r, r[)$ .

*Démonstration.* Soit  $b \in B_d(1, r)$  de la forme  $b = e^{iy}$  avec  $y$  choisi de telle sorte que  $d(1, b) = |y|$ . Vu que  $d(1, b) < r$ , nous avons  $|y| < r$  et donc  $b \in \varphi(]-r, r[)$ .

Dans l'autre sens, si  $y \in ]-r, r[$ , alors

$$d(1, e^{iy}) = \inf_{k \in \mathbb{Z}} |y + 2k\pi| \leq |y| < r. \quad (27.447)$$

EQooRWASooVZnQCJ

Nous avons utilisé le fait que l'infimum sur  $k \in \mathbb{Z}$  est plus petit ou égal à la valeur pour  $k = 0$ . Les inégalités (27.447) montrent que  $e^{iy} \in B_d(1, r)$ . □

Le lemme suivant montre que que les boules autour de 1 sont délimitées par la droite en pointillé de la figure 27.1.

**Lemme 27.111.**

Soit  $r \in [0, \pi]$ . Nous avons

$$B_d(1, r) = S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}. \quad (27.448)$$

*Démonstration.* En deux inclusions.

(1) **Inclusion**  $\subseteq$ 

Si  $r = \pi$ , alors  $B(1, r) = S^1 \setminus \{-1\}$ , alors que  $\cos(\pi) = -1$ .

Si  $r < \pi$ , alors nous partons de la formule (18.18) qui dit que  $e^{ir} = \cos(r) + i \sin(r)$ . D'après le lemme 27.110, un élément de  $B_d(1, r)$  est de la forme  $e^{iy}$  avec  $y \in ]-r, r[$ . Nous voudrions donc prouver que  $\cos(y) > \cos(r)$  dès que  $y \in ]-r, r[$  et  $r < \pi$ .

Sur  $]-r, 0[$ , la fonction  $\cos$  est croissante<sup>52</sup>, donc si  $y < 0$  alors

$$\cos(y) > \cos(-r) = \cos(r). \quad (27.449)$$

De la même façon, sur  $]0, r[$ , la fonction  $\cos$  est décroissante, de telle sorte que si  $y > 0$ , alors  $\cos(y) > \cos(r)$ .

Nous avons prouvé que  $B_d(1, r) \subset S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}$ .

(2) **Inclusion inverse**

Soit  $z = e^{is}$  avec  $s \in [-\pi, \pi[$ , et vérifiant  $\cos(s) > \cos(r)$ . Nous avons

$$d(1, z) = \inf_{k \in \mathbb{Z}} |s + 2k\pi| = |s| \quad (27.450)$$

parce que l'infimum est réalisé par  $k = 0$ . Nous avons donc

$$\begin{cases} \cos(s) > \cos(r) \end{cases} \quad (27.451a)$$

$$\begin{cases} |s| \in [0, \pi] \end{cases} \quad (27.451b)$$

$$\begin{cases} r \in [0, \pi], \end{cases} \quad (27.451c)$$

et nous devons en déduire que  $|s| < r$ . Si  $s > 0$ , alors la décroissance de cosinus (proposition 18.49(1)) dit que  $s < r$ .

Si  $s < 0$ , alors nous savons que  $\cos(s) = \cos(-s)$  et donc, en posant  $s' = -s$ , le cas " $s > 0$ " nous dit que  $s' < r$ , et donc que  $|s| < r$ .

Dans les deux cas nous avons  $d(1, z) = |s| < r$ .

□

**Proposition 27.112.**

La topologie  $\tau_d$  sur  $S^1$ <sup>53</sup> est la topologie induite depuis  $\mathbb{C}$ .

*Démonstration.* Nous notons  $\tau_i$  la topologie induite (c'est-à-dire l'ensemble des ouverts) et  $\tau_d$  la topologie de la distance fraîchement définie. Nous allons également noter  $B_{\mathbb{C}}(a, r)$  la boule dans  $\mathbb{C}$  de centre  $a$  et de rayon  $r$ , et  $B_d(z, r)$  celle dans  $S^1$ , de centre  $z \in S^1$  et de rayon  $r$  pour notre distance  $d$ .

- (1)  $\tau_i \subseteq \tau_d$  Un élément général de  $\tau_i$  est de la forme  $\mathcal{O} \cap S^1$  où  $\mathcal{O}$  est un ouvert de  $\mathbb{C}$ . Soit  $a \in \mathcal{O} \cap S^1$  et prouvons qu'il existe un ouvert de  $\tau_d$  contenant  $a$  et contenu dans  $\mathcal{O} \cap S^1$ ; cela prouvera que  $\mathcal{O} \cap S^1$  est ouvert de  $\tau_d$  par le théorème 7.8.

Soient  $a = e^{ix}$  et  $r$  tel que  $B_{\mathbb{C}}(a, r) \subset \mathcal{O}$ . Nous allons montrer que  $B_d(a, r) \subset B_{\mathbb{C}}(a, r)$ . Un élément général de  $B_d(a, r)$  est  $b = e^{iy}$  tel que

$$d(a, b) = \inf_{k \in \mathbb{Z}} |x - y + 2k\pi| \leq r. \quad (27.452)$$

52. Lemme 18.26.

53. Définition 27.108.

Quitte à redéfinir  $x$  ou  $y$  nous pouvons supposer que l'infimum est atteint en  $k = 0$ . En utilisant la proposition 21.15 nous majorons :

$$|a - b| = |e^{ix} - e^{iy}| \leq |x - y| = d(a, b) \leq r. \quad (27.453)$$

Donc nous avons bien  $b \in B_{\mathbb{C}}(a, r)$  dès que  $b \in B_d(a, r)$ .

(2)  $\tau_d \subset \tau_i$

Ce sens est plus délicat parce que, si nous voulons suivre les mêmes pas que le premier sens, nous devons nous appuyer sur la continuité de l'application  $\ln: \mathbb{C}^* \rightarrow \mathbb{C}$ , laquelle n'est pas vraie en  $-1$  (voir par exemple le lemme 26.34).

Soient  $A \in \tau_d$  et  $a \in A$ . Nous devons prouver l'existence d'un ouvert  $\mathcal{O}$  de  $\mathbb{C}$  tel que  $S^1 \cap \mathcal{O}$  soit inclus dans  $A$  et contienne  $a$ . Nous allons prouver cela dans le cas  $a = 1$  et ensuite propager le résultat en utilisant la symétrie de  $S^1$ .

(2a) Si  $a = 1$  Vu que  $A$  est ouvert pour la topologie de la distance  $d$ , et vu que  $1 \in A$ , il existe  $r > 0$  tel que  $B_d(1, r) \subset A$ . Pour ce  $r$  le lemme 27.111 donne

$$B_d(1, r) = S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}. \quad (27.454)$$

Nous montrons que  $\mathcal{O} = B_{\mathbb{C}}(1, \delta)$  avec  $\delta < 1 - \cos(r)$  fait l'affaire. Si  $x + iy \in B_{\mathbb{C}}(1, \delta)$ , alors  $x > 1 - \delta$  et donc

$$1 - \delta > 1 - (1 - \cos(r)) = \cos(r), \quad (27.455)$$

ce qui prouve que la partie de  $\mathcal{O}$  qui est dans  $S^1$  est bien dans  $B_d(1, r)$ .

(2b) Si  $a \neq 1$

Soient un ouvert quelconque  $A \in \tau_d$  ainsi que  $a = e^{ix} \in A$ . Nous considérons  $r > 0$  tel que  $B_d(a, r) \subset A$ ; nous avons  $e^{-ix}B_d(a, r) \subset e^{-ix}A$  et donc, en tenant compte du lemme 27.109(2) :

$$B_d(1, r) \subset e^{-ix}A. \quad (27.456)$$

Par le premier point, il existe un ouvert  $\mathcal{O}$  de  $\mathbb{C}$  tel que  $1 \in \mathcal{O}$  et

$$\mathcal{O} \cap S^1 \subset B_d(1, r) \subset e^{-ix}A. \quad (27.457)$$

Nous avons évidemment que  $a \in e^{ix}\mathcal{O}$  et

$$e^{ix}(\mathcal{O} \cap S^1) \subset e^{ix}B_d(1, r) \subset A. \quad (27.458)$$

Donc  $e^{ix}\mathcal{O} \cap S^1 \subset A$ . Vu que  $e^{ix}\mathcal{O}$  est un ouvert de  $\mathbb{C}$ , l'ensemble  $e^{ix}\mathcal{O} \cap S^1$  est un ouvert de  $\tau_i$ .

□

LEMooTKFHoJaeMyc

**Lemme 27.113** ([1]).

Deux résultats de limites dans  $S^1$ .

ITEMooEUDIoDuynRg

(1) Pour tout  $a_0 \in S^1$ , nous avons

$$\lim_{s \rightarrow 1} d(a, as) = 0. \quad (27.459)$$

ITEMooXCBUooUxQldB

(2) Si  $f: S^1 \rightarrow \mathbb{C}$  est continue en  $a \in S^1$ , alors

$$\lim_{s \rightarrow 1} f(as) = f(a). \quad (27.460)$$

*Démonstration.* Point par point.

- (1) Soient  $a, s \in S^1$ . Donnons une formule pour  $d(a, as)$ . Si  $a = e^{ix}$  et  $s = e^{iy}$  nous avons  $as = e^{i(x+y)}$  et donc

$$d(a, as) = \inf_{k \in \mathbb{Z}} |x - (x + y) + 2k\pi| = \inf_{k \in \mathbb{Z}} |y + 2k\pi|. \quad (27.461)$$

Voilà pour la formule. Maintenant la preuve de notre point.

Soit  $\epsilon > 0$ . Si  $\delta < \epsilon$  et si  $s \in B(1, \delta)$ , alors il existe  $y \in ]-\delta, \delta[$  tel que  $s = e^{iy}$  par le lemme 27.110. Pour un tel  $s$  nous avons

$$d(a, sa) = \inf_{k \in \mathbb{Z}} |y + 2k\pi| \leq |y| < \delta < \epsilon. \quad (27.462)$$

Nous avons trouvé  $\delta > 0$  tel que  $s \in B(1, \delta)$  implique  $d(a, as) < \epsilon$ . Cela est la limite que nous devons prouver.

- (2) Soit  $\epsilon > 0$ . Soit  $r > 0$  tel que si  $b \in B(a, r)$ , alors  $|g(b) - g(a)| < \epsilon$ ; l'existence d'un tel  $r$  est la continuité de  $g$  en  $a$ . Nous considérons  $\delta > 0$  tel que  $s \in B(1, \delta)$  implique  $sa \in B(a, r)$ ; l'existence d'un tel  $\delta$  est le point (1) de ce lemme.

Avec tout cela nous avons  $|g(as) - g(a)| < \epsilon$  dès que  $s \in B(1, \delta)$ . Nous avons donc, comme nous le voulions, la limite  $\lim_{s \rightarrow 1} g(as) = g(a)$ . □

### Lemme 27.114.

Pour  $s \in S^1$ , nous considérons l'application

$$\begin{aligned} \alpha_s: \text{Fun}(S^1) &\rightarrow \text{Fun}(S^1) \\ \alpha_s(g)(u) &= g(u\bar{s}) - g(u). \end{aligned} \quad (27.463)$$

Quelques propriétés avec  $1 \leq p < \infty$  :

- (1) Si  $f \in L^p(S^1)$ , alors  $\alpha_s(f) \in L^p(S^1)$ .  
 (2) Si  $f$  est continue dans  $L^p(S^1)$  nous avons la limite

$$\lim_{s \rightarrow 1} \alpha_s(f) = 0 \quad (27.464)$$

dans  $L^p(S^1)$ .

*Démonstration.* D'abord un calcul de norme :

$$\|\alpha_s(f)\|_p^p = \int_{S^1} |f(u\bar{s}) - f(u)|^p du \leq \int_{S^1} |f(u\bar{s})|^p du + \int_{S^1} |f(u)|^p du = 2\|f\|_p^p. \quad (27.465)$$

Donc oui pour que  $\alpha_s(f) \in L^p(S^1)$ .

La fonction  $f$  étant supposée continue sur le compact  $S^1$ , elle est majorée. Nous savons qu'en posant  $M = \|f\|_\infty$  nous avons  $|\alpha_s(f)| \leq 2M$ . Donc la fonction constante

$$\begin{aligned} g: S^1 &\rightarrow \mathbb{C} \\ u &\mapsto 2M \end{aligned} \quad (27.466)$$

est une fonction intégrable sur  $S^1$  qui majore  $|\alpha_s(f)|$  uniformément en  $s$ . Soit une suite  $s_i \rightarrow 1$  dans  $S^1$ , et posons  $f_i = \alpha_{s_i}(f)$ . Alors nous avons

$$\|f_i\|_p^p = \int_{S^1} |f_i(u)|^p du \quad (27.467)$$

et aussi  $|f_i|^p \leq (2M)^p$ . Le théorème de la convergence dominée de Lebesgue 14.207 nous permet de permuter limite et intégrale :

$$\lim_{i \rightarrow \infty} \|f_i\|_p^p = \int_{S^1} \lim_{i \rightarrow \infty} |f_i(u)|^p du. \quad (27.468)$$

Mais

$$\lim_{i \rightarrow \infty} f_i(u) = \lim_{i \rightarrow \infty} \alpha_{s_i}(f)(u) = \lim_{i \rightarrow \infty} (f(u\bar{s}_i) - f(u)) = 0. \quad (27.469)$$

La dernière limite est due au fait que  $\lim_{s \rightarrow 1} g(us) = g(u)$  (lemme 27.113(2)). □



### 27.9.3 Système trigonométrique

#### Définition 27.115.

La famille *trigonométrique* sur  $S^1$  est l'ensemble de fonctions  $\{e_n\}_{n \in \mathbb{Z}}$  données par

$$\begin{aligned} e_n : S^1 &\rightarrow \mathbb{C} \\ z &\mapsto z^n \end{aligned} \quad (27.470)$$

avec  $n \in \mathbb{Z}$ . Un *polynôme trigonométrique* est une application  $S^1 \rightarrow \mathbb{C}$  de la forme

$$\sum_{k=-n}^n a_k e_k \quad (27.471)$$

pour des nombres  $a_k \in \mathbb{C}$ , peut-être pas tous non-nuls (autrement dit, il n'est pas forcé d'avoir autant de termes négatifs que positifs).

Le but de  $z \mapsto z^n$  dans cette définition est d'être lu  $t \mapsto e^{int}$  lorsqu'on considère les fonctions sur  $[0, 2\pi[$ .

PROPOOMGFooROFFFr

#### Proposition 27.116.

La famille trigonométrique est une famille orthonormale pour le produit scalaire  $L^2(S^1, \mathcal{A}, \mu)$ .

*Démonstration.* En utilisant la proposition 27.107(3) nous avons :

$$\langle e_n, e_n \rangle = \int_{S^1} e_n \overline{e_n} = \frac{1}{2\pi} \int_{[0, 2\pi[} e_n(\varphi(x)) \overline{e_n(\varphi(x))} \quad (27.472a)$$

$$= \frac{1}{2\pi} \int_{[0, 2\pi[} e^{inx} e^{-inx} dx \quad (27.472b)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} 1 dx \quad (27.472c)$$

$$= 1. \quad (27.472d)$$

Et nous avons également, pour  $m \neq n$  :

$$\langle e_n, e_m \rangle = \frac{1}{2\pi} \int_{[0, 2\pi[} e^{i(n-m)x} dx = \frac{1}{2\pi} \left[ \frac{1}{i(n-m)} e^{i(n-m)x} \right]_0^{2\pi} = 0. \quad (27.473)$$

□

#### Remarque 27.117.

Vous aurez noté que le facteur  $\frac{1}{2\pi}$  qui permet d'avoir  $\langle e_n, e_n \rangle = 1$  ne provient ni de la définition du produit scalaire ni de celle de la famille trigonométrique, mais bien de la mesure, voir la définition 18.194.

REMooUCANooVyXPxj

#### Remarque 27.118.

Notez aussi que nous avons bien  $\langle e_n, e_{-n} \rangle = 0$ . Il faut donc bien prendre tous les  $e_n$  avec  $n \in \mathbb{Z}$  et non seulement  $n \in \mathbb{N}$ .

PROPOoTGBHooXGhdPR

#### Proposition 27.119.

Les polynômes trigonométriques forment une partie dense dans  $(C(S^1, \mathbb{C}), \|\cdot\|_\infty)$ .

*Démonstration.* Pour préciser les notations,  $C(S^1, \mathbb{C})$  est l'ensemble des fonctions continues de  $S^1$  vers  $\mathbb{C}$ , et l'espace topologique que nous considérons est cet ensemble sur lequel nous considérons la distance supremum.

Nous utilisons le théorème de Stone-Weierstrass 12.444.

Le système contient une fonction constante non nulle, à savoir  $e_0$ .

Il sépare les points grâce à la fonction  $e_1$  qui n'est autre que la fonction identité  $z \mapsto z$ . De plus l'ensemble des polynômes trigonométriques est stable par conjugaison parce que si

$$P = \sum_{k=-n}^n a_k e_k, \quad (27.474)$$

alors  $\bar{P} = \sum_{k=-n}^n \overline{a_k e_k} = \sum_{k=-n}^n \overline{a_k} e_{-k}$  qui est encore un polynôme trigonométrique.  $\square$

### Définition 27.120.

Si nous avons une fonction  $f: S^1 \rightarrow \mathbb{C}$ , nous définissons ses **coefficients de Fourier** par

$$c_n(f) = \langle f, e_n \rangle \quad (27.475)$$

pourvu que l'intégrale existe.

## 27.9.4 Convolution

La convolution sur  $\mathbb{R}^n$  est donnée par la définition 27.78. Nous voyons maintenant comment cela s'adapte à  $S^1$ .

### Définition 27.121.

Si  $f$  et  $g$  sont des fonctions sur  $S^1$  à valeurs dans  $\mathbb{C}$ , nous définissons la **convolution** de  $f$  et  $g$  comme étant la fonction sur  $S^1$  définie par

$$(f * g)(z) = \int_{S^1} f(s)g(z\bar{s})d\mu(s). \quad (27.476)$$

Cette définition appelle plusieurs remarques.

- Dès que  $z, s \in S^1$ , nous avons  $z\bar{s} \in S^1$ , de telle sorte qu'au moins l'intégrande ait un sens.
- Nous ne prétendons pas que l'intégrale (27.476) converge pour toutes les fonctions  $f$  et  $g$ . Cela est une définition « pour tous les couples  $f, g$  pour lesquels l'intégrale fonctionne ».
- Le lemme 27.122 nous dira que  $L^1(S^1)$  est stable par convolution : si  $f$  et  $g$  sont dans  $L^1$ , alors  $f * g$  y est aussi.
- Dans la formule (27.476), la variable  $s$  est vraiment une variable muette. Cette formule aurait également pu être écrite

$$(f * g)(z) = \int_{S^1} [s \mapsto f(s)g(z\bar{s})]d\mu. \quad (27.477)$$

### Lemme 27.122 ([748]).

Si  $f, g \in L^1(S^1)$ , alors pour presque tout  $z \in S^1$ , la fonction  $s \mapsto f(s)g(z\bar{s})$  est dans  $L^1(S^1)$ .

*Démonstration.* Nous considérons la fonction

$$\begin{aligned} \psi: S^1 \times S^1 &\rightarrow \mathbb{C} \\ (z, s) &\mapsto f(s)g(z\bar{s}). \end{aligned} \quad (27.478)$$

- (1)  $\psi \in L^1(S^1 \times S^1)$  Nous utilisons le corolaire 14.304, et pour cela nous calculons les intégrales en chaîne<sup>54</sup> :

$$\int_{S^1} \left[ \int_{S^1} |f(s)g(z\bar{s})|dz \right] ds = \int_{S^1} |f(s)| \underbrace{\left[ \int_{S^1} |g(z\bar{s})|dz \right]}_{=A < \infty} ds \quad (27.479a)$$

$$= A \int_{S^1} |f(s)|ds \quad (27.479b)$$

$$< \infty. \quad (27.479c)$$

<sup>54</sup>. Dans les expressions suivantes, les symboles «  $ds$  » et «  $dz$  » n'ont pas d'autres valeurs que purement de notation pour indiquer le nom de la variable d'intégration.

Le fait que  $A < \infty$  provient directement de l'hypothèse  $g \in L^1(S^1)$  <sup>55</sup>.

Par le corolaire sus-cité nous avons bien  $\psi \in L^1(S^1 \times S^1)$ .

- (2) **Et par Fubini** Le théorème de Fubini 14.305(1) nous renseigne que pour presque tout  $z \in S^1$ , l'application

$$s \mapsto \psi(z, s) \quad (27.480)$$

est dans  $L^1(S^1)$ . Et la partie 14.305(2) ajoute que l'application

$$z \mapsto \int_{S^1} \psi(s, z) ds \quad (27.481)$$

est également  $L^1(S^1)$ .

- (3) **Conclusion** L'application donnée en (27.481) est précisément  $(f * g)$ . Donc  $f * g \in L^1(S^1)$ . □

**Lemme 27.123** ([1]).

Si  $f \in L^1(S^1)$  et si  $g$  est continue sur  $S^1$ , alors  $f * g$  existe et est continue sur  $S^1$ .

*Démonstration.* Vu que  $S^1$  est compact, la continuité de  $g$  implique que  $g$  est bornée et donc dans  $L^1(S^1)$ . Le lemme 27.122 dit alors que  $f * g$  est bien définie sur  $S^1$ .

Soit  $z_0 \in S^1$ . Nous montrons que  $f * g$  est continue en  $z_0$ ; pour cela nous considérons  $\epsilon > 0$  et ensuite nous réfléchissons un peu.

Vu que  $g$  est continue sur  $S^1$  qui est compact,  $g$  y est uniformément continue par le théorème de Heine 12.80. Il existe donc un  $\delta > 0$  tel que pour tout  $z_0 \in S^1$ , si  $z \in B(z_0, \delta)$ , alors  $|g(z_0) - g(z)| < \epsilon$ .

Soit  $s \in S^1$ . Si  $z \in B(z_0, \delta)$ , alors  $\bar{s}z \in B(\bar{s}z_0, \delta)$  par le lemme 27.109(2). Dans ce cas nous avons aussi

$$|g(\bar{s}z_0) - g(\bar{s}z)| < \epsilon. \quad (27.482)$$

Un peu de calcul maintenant. D'une part

$$(f * g)(z_0) - (f * g)(z) = \int_{S^1} f(s)(g(z_0\bar{s}) - g(z\bar{s}))ds, \quad (27.483)$$

et donc

$$|(f * g)(z_0) - (f * g)(z)| \leq \int_{S^1} |f(s)| \underbrace{|g(z_0\bar{s}) - g(z\bar{s})|}_{< \epsilon} ds \quad (27.484a)$$

$$\leq \epsilon \int_{S^1} |f| \quad (27.484b)$$

$$= A\epsilon \quad (27.484c)$$

pour une certaine constante  $A$  ne dépendant pas de  $z_0$ .

Nous avons prouvé que pour tout  $\epsilon > 0$ , il existe un  $\delta$  tel que  $z \in B(z_0, \alpha)$  implique

$$|(f * g)(z_0) - (f * g)(z)| \leq A\epsilon, \quad (27.485)$$

ce qui signifie que  $f * g$  est continue en  $z_0$ . □

Notez que dans cette démonstration, l'uniforme continuité de  $g$  a été utilisée pour effectuer d'un seul coup la majoration pour tout  $s$  dans l'intégrale.

**Proposition 27.124.**

Si  $f \in L^1(S^1)$ , nous avons

$$f * e_n = c_n(f)e_n. \quad (27.486)$$

---

55. Avec un changement de variables  $z \mapsto z\bar{s}$  que je vous conseille d'être capable de justifier.

*Démonstration.* Il s'agit d'un bon calcul. En considérant  $z = e^{i\theta}$  nous avons

$$(f * e_n)(z) = \int_{S^1} f(s) e_n(z\bar{s}) ds \quad (27.487a)$$

$$= \frac{1}{2\pi} \int_{[0, 2\pi[} f(e^{ix}) e_n(e^{i(\theta-x)}) dx \quad (27.487b)$$

$$= e^{in\theta} \frac{1}{2\pi} \int_{[0, 2\pi[} f(e^{ix}) e^{-inx} dx \quad (27.487c)$$

$$= e_n(e^{i\theta}) \frac{1}{2\pi} \int_{[0, 2\pi[} f(e^{ix}) \overline{e_n(e^{ix})} dx \quad (27.487d)$$

$$= e_n(z) \int_{S^1} f(s) \overline{e_n(s)} ds \quad (27.487e)$$

$$= e_n(z) \langle f, e_n \rangle. \quad (27.487f)$$

Donc  $(f * e_n)(z) = \langle f, e_n \rangle e_n(z)$ , c'est-à-dire que

$$f * e_n = c_n(f) e_n. \quad (27.488)$$

□

LEMooDGHJooRAnwpy

### Lemme 27.125.

Si  $P$  est un polynôme trigonométrique et si  $f \in L^1(S^1)$ , alors  $f * P$  est également un polynôme trigonométrique.

*Démonstration.* Soit  $P = \sum_{k=-n}^n a_k e_k$ . Par la linéarité du produit de convolution,

$$f * P = \sum_{k=-n}^n a_k f * e_k = \sum_k a_k c_k(f) e_k \quad (27.489)$$

où nous avons également utilisé la proposition 27.124. Nous avons donc un polynôme trigonométrique dont les coefficients sont  $a_k c_k(f)$  au lieu de  $a_k$ . □

## 27.9.5 Approximation de l'unité

LEMooUNFBooRCzwIn

### Lemme 27.126 ([724]).

Soient une fonction continue  $f: S^1 \rightarrow [0, \infty[$  et  $a \in S^1$  telle que  $f(z) < f(a)$  pour tout  $z \in S^1 \setminus \{a\}$ . Alors la suite de fonctions  $f_n: S^1 \rightarrow \mathbb{R}$  donnée par

$$f_n(z) = \left( \int_{S^1} f^n \right)^{-1} f(z)^n \quad (27.490)$$

EQooTQYPooLZprJi

est une approximation de l'unité<sup>56</sup> autour de  $a$ .

*Démonstration.* En plusieurs points, dont d'abord une série de vérifications pour voir que la formule a un sens.

- (1) **Strictement positif** D'abord, vu que  $f$  prend ses valeurs dans  $[0, \infty[$  et vu que  $f(z) < f(a)$ , nous avons  $f(a) > 0$  (strict). Peut-être que  $f$  s'annule à certains endroits de  $S^1$ , mais pas  $a$ .
- (2)  **$f^n$  est intégrable sur  $S^1$**  La fonction  $f^n$  est dans les hypothèses de la proposition 14.218 parce que  $S^1$  est compact,  $f^n$  y est continue et la mesure sur  $S^1$  est compatible avec la topologie (voir les hypothèses précises).

---

56. Définition 27.90.

- (3) **L'intégrale n'est pas nulle** Vu que  $f(a) > 0$ , il existe un ouvert  $A$  contenant  $a$  sur lequel  $f > 0$ . Nous avons alors

$$\int_K f^n \geq \int_A f^n > 0. \quad (27.491)$$

Cela pour dire que l'inverse dans (27.490) ne pose pas de problèmes.

- (4) **Norme** Vu que toutes les fonctions tant  $f$  que  $f_n$  sont positives, les valeurs absolues ne jouent aucun rôle et nous avons

$$\|f_n\|_1 = \int_{S^1} f_n d\mu = \left( \int_{S^1} f^n \right)^{-1} \int_{S^1} |f(z)|^n d\mu = 1. \quad (27.492)$$

Ce calcul donne d'un seul coup les deux conditions

- $\sup_k \|f_k\| = 1$
- $\int_{S^1} f_n = 1$  pour tout  $n$ .

Nous passons maintenant au vrai travail. Soit un voisinage  $V$  de  $a$  dans  $S^1$ . Soit une suite croissante  $(t_k)$  qui converge vers  $f(a)$ , c'est-à-dire  $0 < t_k < f(a)$ . Nous posons

$$A_k = \{x \in S^1 \setminus V \text{ tel que } f(x) \geq t_k\}. \quad (27.493)$$

Cet ensemble est contenu dans  $S^1$  et est donc borné (pour la métrique de  $S^1$ ).

- (1)  **$A_k$  est fermé** Attention : ici nous démontrons que  $A_k$  est fermé dans  $S^1$ , et les complémentaires sont pris dans  $S^1$ .

Nous montrons que le complémentaire est ouvert en prenant  $y \in A^c$  et en montrant que  $y$  admet un voisinage contenu dans  $A^c$  (le fameux théorème 7.8 que nous ne nous lasserons jamais de citer). Si  $y \in A^c$ , il y a deux possibilités (non exclusives) : soit  $y \in V$  soit  $f(y) < t_k$ . Si  $y \in V$ , alors le voisinage  $V$  lui-même est encore dans  $A^c$ . Si par contre  $f(y) < t_k$ , alors par continuité, il existe un voisinage de  $y$  sur lequel  $f < t_k$ .

- (2)  **$A_k$  est compact** L'espace  $S^1$  est compact, par exemple grâce au lemme 7.112. La partie  $A_k$  est fermée dans le compact  $S^1$ , donc elle est compacte par le lemme 7.121.
- (3)  **$A_{k+1} \subset A_k$**  Si  $x \in A_{k+1}$ , alors  $f(x) \geq t_{k+1} > t_k$ . Donc  $f(x) > t_k$  et  $x \in A_k$ .
- (4) **Intersection vide** Si  $x \in \bigcap_{k \in \mathbb{N}} A_k$ , alors  $f(x) \geq t_k$  pour tout  $k$ . En passant à la limite et en sachant que  $\lim_{k \rightarrow \infty} t_k = f(a)$ , nous avons  $f(x) \geq f(a)$ . Par hypothèse, cela n'est pas. Donc

$$\bigcap_{k \in \mathbb{N}} A_k = \emptyset. \quad (27.494)$$

Nous avons, dans un compact, des fermés emboîtés dont l'intersection est vide. Le corolaire 7.118 nous dit qu'il existe un indice à partir duquel tous les  $A_k$  sont vides.

Soit  $\delta = t_k$  pour un  $k$  tel que  $A_k$  est vide. Nous avons

$$\{x \in S^1 \setminus V \text{ tel que } f(x) \geq \delta\} = \emptyset, \quad (27.495)$$

c'est-à-dire que sur  $S^1 \setminus V$ , nous avons  $f < \delta$  et donc

$$\int_{S^1 \setminus V} f(s)^n < \int_{S^1 \setminus V} \delta^n = \text{Vol}(S^1 \setminus V) \delta^n \quad (27.496)$$

où  $\text{Vol}(S^1 \setminus V) = \int_{S^1 \setminus V} 1 = \mu'(S^1 \setminus V)$  est une constante réelle strictement positive.

Nous avons aussi  $\delta < f(a)$  parce que  $\delta$  est un des  $t_k$  (et que cette suite croissante converge vers  $f(a)$  sans l'atteindre par hypothèse). Soit  $\delta_1$  tel que  $\delta < \delta_1 < f(a)$ <sup>57</sup>.

Nous posons

$$W = \{x \in S^1 \text{ tel que } f(x) > \delta_1\}. \quad (27.497)$$

57. Dans [724], il prend  $\delta < \delta_1 < 1$  et je crois qu'il aurait dû écrire  $\varphi(0)$  au lieu de 1.

Cet ensemble n'est pas vide parce qu'il contient  $a$  et est ouvert parce que  $f$  est continue. Nous avons

$$\int_{S^1} f(s)^n ds \geq \int_W f(s)^n ds \geq \delta_1^n \text{Vol}(W). \quad (27.498)$$

Nous avons donc déjà ces deux inégalités :

$$\int_{S^1 \setminus V} f(s)^n \leq \text{Vol}(S^1 \setminus V) \delta^n \quad (27.499)$$

et

$$\int_{S^1} f(s)^n ds \geq \delta_1^n \text{Vol}(W). \quad (27.500)$$

En ce qui concerne les fonctions  $f_n$  que nous voulions étudier,

$$f_n(z) = \left( \int_{S^1} f(s)^n ds \right)^{-1} f(z)^n \leq (\text{Vol}(W) \delta_1^n)^{-1} f(z)^n, \quad (27.501)$$

et donc

$$\int_{S^1 \setminus V} f_n \leq (\text{Vol}(W) \delta_1^n)^{-1} \text{Vol}(S^1 \setminus V) \sigma^n = \frac{\text{Vol}(S^1 \setminus V)}{\text{Vol}(W)} \left( \frac{\delta}{\delta_1} \right)^n. \quad (27.502)$$

Étant donné que  $\delta < \delta_1$ , nous avons  $(\delta/\delta_1)^n \rightarrow 0$ . Donc aussi

$$\lim_{n \rightarrow \infty} \int_{S^1 \setminus V} f_n(z) dz = 0. \quad (27.503)$$

□

Le théorème suivant est une version pour  $S^1$  du théorème 27.93. Le produit de convolution dans  $S^1$  est la définition 27.121.

THOooIAOPooELSNxq

**Théorème 27.127** ([724, 1]).

Soient  $(\varphi_k)$  une approximation de l'unité sur  $\Omega = S^1$  ainsi qu'une fonction  $g: S^1 \rightarrow \mathbb{C}$ . Si  $g$  est mesurable et bornée sur  $\Omega$  et si  $g$  est continue en  $a_0$  alors

$$(\varphi_k * g)(a_0) \rightarrow g(a_0). \quad (27.504)$$

*Démonstration.* Pour chaque  $k \in \mathbb{N}$  nous posons

$$d_k = \varphi_k * g - g. \quad (27.505)$$

Le but de ce théorème est de montrer que  $d_k \rightarrow 0$  pour diverses notions de convergence.

Soit  $a_0 \in S^1$ . Par définition de l'approximation de l'unité,  $\int_{S^1} \varphi_k = 1$  et donc on peut écrire  $g(a_0) = \int_{S^1} g(a_0) \varphi_k(s) ds$ . En ce qui concerne  $d_k(a_0)$  nous avons alors

$$d_k(a_0) = \int_{S^1} \varphi_k(s) g(a_0 \bar{s}) ds - \int_{S^1} g(a_0) \varphi_k(s) ds \quad (27.506a)$$

$$= \int_{S^1} \varphi_k(s) (g(a_0 \bar{s}) - g(a_0)) ds. \quad (27.506b)$$

Nous pouvons passer à la norme (et non la valeur absolue parce que  $d_k$  prend ses valeurs dans  $\mathbb{C}$ ) :

$$|d_k(a_0)| \leq \int_{S^1} |\varphi_k(s)| |g(a_0 \bar{s}) - g(a_0)| ds. \quad (27.507)$$

La définition d'une approximation de l'unité nous permet de considérer  $M = \sup_k \|\varphi_k\|_1 < \infty$ . Le lemme 27.113(2) nous permet, lui, de considérer  $\alpha > 0$  tel que

$$|g(a_0 \bar{s}) - g(a_0)| < \epsilon \quad (27.508)$$

dès que  $s \in B(1, \alpha)$ <sup>58</sup>. Vu que la suite  $(\varphi_k)$  est une approximation de l'unité, nous avons

$$\lim_{k \rightarrow \infty} \int_{S^1 \setminus B(1, \alpha)} |\varphi_k| = 0. \quad (27.509)$$

Soit  $k$  suffisamment grand pour avoir  $\int_{S^1 \setminus B(1, \alpha)} |\varphi_k| < \epsilon$ . Avec tout cela nous avons les majorations

$$|d_k(a_0)| \leq \int_{S^1} |\varphi_k(s)| |g(a_0 \bar{s}) - g(a_0)| ds \quad (27.510a)$$

$$= \int_{S^1 \setminus B(1, \alpha)} |\varphi_k(s)| \underbrace{|g(a_0 \bar{s}) - g(a_0)|}_{\leq 2\|g\|_\infty} ds + \int_{B(1, \alpha)} |\varphi_k(s)| \underbrace{|g(a_0 \bar{s}) - g(a_0)|}_{< \epsilon} ds \quad (27.510b)$$

$$\leq 2\|g\|_\infty \int_{S^1 \setminus B(1, \alpha)} |\varphi_k(s)| ds + \epsilon \int_{S^1} |\varphi_k(s)| ds \quad (27.510c)$$

$$\leq \epsilon(1 + 2\|g\|_\infty). \quad (27.510d)$$

Nous avons donc bien  $\lim_{k \rightarrow \infty} |d_k(a_0)| = 0$  et donc la continuité de  $\varphi_k * g$  en  $a_0$ . □

Voici une version un peu forte sous l'hypothèse de continuité. Vu que  $S^1$  est compact, la continuité est en réalité une hypothèse assez forte : ça implique l'uniforme continuité et l'existence d'un maximum et d'un minimum.

**Proposition 27.128.**

Soient  $(\varphi_k)$  une approximation de l'unité sur  $\Omega = S^1$  ainsi qu'une fonction continue  $g: S^1 \rightarrow \mathbb{C}$ .

(1) Si  $g \in L^p(\Omega)$  ( $0 \leq p < \infty$ ) et si  $g$  est continue, alors<sup>59</sup>

$$\varphi_k * g \xrightarrow{L^p} g. \quad (27.511)$$

(2) Si  $g$  est continue sur  $S^1$ , alors

$$\varphi_k * g \xrightarrow{L^\infty} g \quad (27.512)$$

*Démonstration.* En plusieurs points

(1)  $\lim_{u \rightarrow 1} \|\tau_u(g) - g\| = 0$  Ceci est un petit point intermédiaire. Pour des besoins de notations, nous posons

$$\begin{aligned} \tau_u(g): S^1 &\rightarrow \mathbb{C} \\ s &\mapsto g(s\bar{u}) \end{aligned} \quad (27.513)$$

pour  $u \in S^1$ .

La fonction  $g$  est continue sur le compact  $S^1$ , et y est donc uniformément continue<sup>60</sup>. Nous allons en déduire que  $\lim_{u \rightarrow 1} \|\tau_u(g) - g\|_\infty = 0$ .

Soit  $\epsilon > 0$ . L'uniforme continuité de  $g$  signifie qu'il existe  $\delta > 0$  tel que pour tout  $a \in S^1$  si  $s \in B(a, \delta)$ , alors  $|g(s) - g(a)| < \epsilon$ . Si  $u \in B(1, \delta)$  nous avons aussi  $\bar{u} \in B(1, \delta)$  et donc  $s\bar{u} \in B(s, \delta)$ ; ça c'est le lemme 27.109(2).

Pour tout  $a \in S^1$  nous avons la chaîne

$$u \in B(1, \delta) \Rightarrow a\bar{u} \in B(a, \delta) \Rightarrow |g(a\bar{u}) - g(a)| < \epsilon. \quad (27.514)$$

Cela étant valable pour tout  $a$ , c'est encore valable en passant au supremum<sup>61</sup> :

$$u \in B(1, \delta) \Rightarrow \sup_{a \in S^1} |g(a\bar{u}) - g(a)| \leq \epsilon \quad (27.515)$$

58. Notez que  $s \in B(1, \alpha)$  si et seulement si  $\bar{s} \in B(1, \alpha)$ . Il n'y a donc pas d'incohérence entre l'hypothèse sur  $s$  et notre condition sur  $g(a_0 \bar{s})$

59. Vous noterez les  $p \in ]0, 1[$  en bonus par rapport au cas de  $\mathbb{R}^n$ .

60. Théorème de Heine 12.80. C'est fondamentalement ce fait qui unifie les parties (1) et (2) de cette preuve.

61. Notez l'inégalité qui n'est plus stricte.

et donc d'accord pour

$$\lim_{u \rightarrow 1} \|\tau_u(g) - g\| = 0. \quad (27.516)$$

- (2)  $\|d_k\|_\infty \rightarrow 0$  Nous prouvons la convergence uniforme sur  $S^1$  de  $d_k$  vers zéro. Ensuite nous verrons que la compacité de  $S^1$  permet d'en déduire les points (1) et (2).

En utilisant la notation  $\tau_u$ , nous pouvons écrire

$$d_k(s) = \int_{S^1} \varphi_k(u) g(s\bar{u}) du - g(s) = \int_{S^1} \varphi_k(u) \left( \underbrace{g(s\bar{u})}_{=\tau_u(g)(s)} - g(s) \right) du, \quad (27.517)$$

et donc

$$|d_k(s)| \leq \int_{S^1} |\varphi_k(u)| \|\tau_u(g) - g\|_\infty du. \quad (27.518)$$

Nous posons  $M = \sup_k \|\varphi_k\|_1$ , et nous considérons  $\delta$  tel que  $\|\tau_u(g) - g\|_\infty < \epsilon$  pour tout  $u \in B(1, \delta)$ . Ensuite nous subdivisons  $S^1$  en  $B(1, \delta)$  et  $B(1, \delta)^c$  :

$$\|d_k\|_\infty \leq \int_{B(1, \delta)} |\varphi_k(u)| \underbrace{\|\tau_u(g) - g\|_\infty}_{\leq \epsilon} du + \int_{B(1, \delta)^c} |\varphi_k(u)| \|\tau_u(g) - g\|_\infty du \quad (27.519a)$$

$$\leq \epsilon M + 2\|g\|_\infty \int_{B(1, \delta)^c} |\varphi_k(u)| du \quad (27.519b)$$

$$\leq \epsilon(M + 2\|g\|_\infty) \quad (27.519c)$$

parce que pour chaque  $s \in S^1$  nous avons  $\tau_u(g)(s) - g(s)$  et donc  $\|\tau_u(g) - g\|_\infty \leq 2\|g\|_\infty$ .

Tout cela montre que  $d_k \xrightarrow{\|\cdot\|_\infty} 0$ .

- (3) Convergence  $L^p$ ,  $0 < p < \infty$

Soit  $\epsilon > 0$ . Nous avons

$$\|d_k\|_p^p \leq \int_{S^1} |(\varphi_k * g)(s) - g(s)|^p ds. \quad (27.520)$$

Il existe un  $k$  à partir duquel  $\|\varphi_k * g - g\|_\infty < \epsilon$ . Pour de tels  $k$  nous avons

$$\|d_k\|_p^p < \epsilon^p. \quad (27.521)$$

Ce passage est très possible dans le cas de  $S^1$  parce que  $\int_{S^1} 1 = 1$ . Dans le cas de  $\mathbb{R}^d$ , c'est pas du tout bon ; c'est pour cela que nous avons un résultat un peu plus fort dans  $S^1$ . La croissance de la fonction puissance (proposition 12.427) nous permet de conclure que  $\|d_k\|_p < \epsilon$ .

Nous avons donc la convergence  $L^p$  pour  $0 < p < \infty$ .

- (4) Convergence  $L^\infty$

Non, la convergence  $L^\infty$  n'est pas la convergence pour la norme  $\|\cdot\|_\infty$ . Voir la sous-section 27.3.3. Il n'en reste pas moins que si  $\epsilon > 0$  et si  $k$  est assez grand pour que  $\|f_k - f\|_\infty < \epsilon$ , nous aurons

$$N_\infty(f_k - f) \leq \|f_k - f\| < \epsilon. \quad (27.522)$$

□

### 27.9.6 Base hilbertienne (suite des polynômes trigonométriques)

Voici le plan pour la suite :

- Construire un polynôme trigonométrique qui vérifie les hypothèses du lemme 27.126.
- En déduire une approximation de l'unité constituée de polynômes trigonométriques.
- Dire que si  $f \in L^2(S^1)$ , alors  $f * \varphi_k$  est un polynôme trigonométrique dès que  $\varphi_k$  en est un.



— Invoquer le théorème 27.127(1) pour déduire que  $\varphi_k * f$  est une suite de polynômes trigonométriques dans  $L^2(S^1)$  qui converge  $\varphi_k * f \xrightarrow{L^2} f$ .

LEMooQQILooWlhntZ

**Lemme 27.129.**

La fonction  $P = e_1 + e_{-1}$  est continue à valeurs réelles sur  $S^1$ .

*Démonstration.* Nous avons  $e_1(z) = z$  et  $e_{-1}(z) = z^{-1}$ , c'est-à-dire que pour  $z = e^{ix}$  ( $x \in \mathbb{R}$ ), nous avons  $e_{-1}(e^{ix}) = e^{-ix}$ , de telle sorte que, en utilisant le lemme 18.11 qui donne  $e^{ix}$  en termes des fonctions trigonométriques usuelles :

$$(e_1 + e_{-1})(e^{ix}) = e^{ix} + e^{-ix} = \cos(x) + i \sin(x) + \cos(x) - i \sin(x) = 2 \cos(x). \quad (27.523)$$

Nous avons donc la continuité et les valeurs réelles. □

LEMooIDTVooYTpfEm

**Lemme 27.130.**

Il existe un polynôme trigonométrique à valeurs dans  $[0, \infty[$  et tel que  $f(a) < f(1)$  pour tout  $a \neq 1$  dans  $S^1$ .

*Démonstration.* Le lemme 27.129 nous dit déjà que  $P = e_1 + e_{-1}$  est continue à valeurs réelles. Or qui est continue sur un compact (ici  $S^1$ ), atteint donc ses bornes. Il est donc facile de considérer<sup>62</sup>  $M > 0$  tel que  $Q = M + e_1 + e_{-1}$  est à valeurs dans  $[0, \infty[$ .

Une forme explicite de  $Q$  est que

$$Q(e^{ix}) = M + 2 \cos(x). \quad (27.524)$$

Le maximum de  $\cos(x)$  est obtenu en  $x = 0$  et vaut 1. Le maximum de  $Q$  est alors  $Q(1) = 2 + M$ . Il n'est atteint qu'une seule fois sur  $S^1$  parce que pour avoir  $Q(e^{ix}) = 2 + M$ , il faut avoir  $2 \cos(x) = 2$ , c'est-à-dire  $x = 2k\pi$ . Mais  $e^{i2k\pi} = 1$ .

Donc  $Q(a) < M + 2 = Q(1)$  pour tout  $a \neq 1$  dans  $S^1$ . □

**Proposition 27.131.**

Les polynômes trigonométriques  $\{e_n\}_{n \in \mathbb{Z}}$  forment une base hilbertienne de  $L^2(S^1)$ .

*Démonstration.* Le fait que les  $e_n$  soient orthonormée est la proposition 27.116. Il reste à prouver que ce soit un système total.

Soit  $f \in L^2(S^1)$ . Soit un polynôme  $Q$  vérifiant le lemme 27.130 ; nous posons

$$\varphi_k(z) = \left( \int_{S^1} Q^n \right)^{-1} Q(z)^n. \quad (27.525)$$

Cela est une approximation de l'unité par la proposition 27.126. Les  $\varphi_k$  sont des polynômes trigonométriques parce que les  $Q^n$  le sont et que  $\int_{S^1} Q^n$  est seulement un nombre.

Le lemme 27.125 nous dit alors que pour tout  $k$ , la fonction

$$\varphi_k * f \quad (27.526)$$

est un polynôme trigonométrique. □

Nous nous permettons de confirmer la remarque 27.118 comme quoi il faut bien tous les  $e_n$  avec  $n \in \mathbb{Z}$ , parce que le polynôme trigonométrique  $Q$  est bien construit à partir de  $e_1 + e_{-1}$ .

62. Par exemple,  $M$  est le maximum de  $|P|$ .

### 27.9.7 Convolution, bis

LEMooLUBQooWLMFrN

#### Lemme 27.132.

Nous considérons l'application

$$\begin{aligned}\varphi: \mathbb{R} &\rightarrow S^1 \\ t &\mapsto e^{it}.\end{aligned}\tag{27.527}$$

Soient  $t, u \in \mathbb{R}$  tels que  $\varphi(t) = \varphi(u)$ . Alors pour toutes fonctions pour lesquelles les intégrales convergent,

$$\int_0^{2\pi} (f \circ \varphi)(\theta)(g \circ \varphi)(t - \theta) \frac{d\theta}{2\pi} = \int_0^{2\pi} (f \circ \varphi)(\theta)(g \circ \varphi)(u - \theta) \frac{d\theta}{2\pi}.\tag{27.528}$$

*Démonstration.* Si  $\varphi(u) = \varphi(t)$ , alors  $u = t + 2k\pi$  pour un certain  $k \in \mathbb{Z}$ . Cette condition implique que  $\varphi(t - \theta) = \varphi(u - \theta)$ , et donc l'égalité

$$\int_0^{2\pi} (f \circ \varphi)(\theta)(g \circ \varphi)(t - \theta) \frac{d\theta}{2\pi} = \int_0^{2\pi} (f \circ \varphi)(\theta)(g \circ \varphi)(u - \theta) \frac{d\theta}{2\pi}.\tag{27.529}$$

□

#### Définition 27.133 (Convolution sur $S^1$ ).

Le lemme 27.132 permet de définir

$$(f * g)(\varphi(t)) = \int_0^{2\pi} (f \circ \varphi)(\theta)(g \circ \varphi)(t - \theta) \frac{d\theta}{2\pi}\tag{27.530}$$

pour toutes les paires de fonctions  $f, g \in \text{Fun}(S^1, \mathbb{C})$  pour lesquelles l'intégrale converge.

L'espace de Lebesgues  $L^2([a, b])$  sera traité dans 29.6.

## 27.10 Sur $[0, 2\pi[$

Le produit de convolution est un peu subtil parce que  $f(t - x)$  n'est pas défini à priori pour tout  $t, x \in [0, 2\pi[$ , vu que  $f$  n'est définie que sur  $[0, 2\pi[$ . Au moins trois solutions s'offrent à nous :

- considérer implicitement la fonction prolongée par périodicité.
- considérer les fonctions sur  $\mathbb{R}/2\pi$ , et définir un peu toutes les opérations modulo  $2\pi$  (fastidieux)
- utiliser une bijection ayant les bonnes propriétés avec  $S^1$  sur lequel tout est déjà fait.

Nous sélectionnons la troisième voie. Pour cela nous considérons la fonction (attention, elle n'est pas tout à fait la même que celle plus haut)

$$\begin{aligned}\varphi: [0, 2\pi[ &\rightarrow S^1 \\ t &\mapsto e^{it}\end{aligned}\tag{27.531}$$

qui est une bijection par la proposition 18.62(1). Pour le produit de convolution,

$$(f * g)(x) = (f \circ \varphi^{-1}) * (g \circ \varphi^{-1})(\varphi(x))\tag{27.532}$$

pour toutes les fonctions  $f, g: [0, 2\pi[ \rightarrow \mathbb{C}$  pour lesquelles les intégrales en jeu ont un sens.

## 27.11 Des inégalités

Avant d'entrer dans le vif du sujet, nous nous fendons d'une petite étude de fonction. Soit

$$\begin{aligned}\phi: [0, 1] &\rightarrow \mathbb{R} \\ x &\mapsto \frac{(1+x)^r}{1+x^r}.\end{aligned}\tag{27.533}$$

Un peu de calcul montre que

$$\frac{\phi'(x)}{\phi(x)} = \frac{r(1-x^{r-1})}{(1+x^r)(1+x)}.\tag{27.534}$$

### 27.11.1 c-r inégalité

LEMooFKKEooDTypUd

#### Lemme 27.134.

Soient  $a, b > 0$  et  $r \geq 1$ . Nous avons les inégalités

$$a^r + b^r \leq (a + b)^r \leq 2^{r-1}(a^r + b^r). \quad (27.535)$$

*Démonstration.* Notons d'abord que si  $r = 1$ , nous avons juste de égalités ; pas de problèmes. Nous supposons donc que  $r > 1$ . Pour la première inégalité, nous posons  $f(x) = a^r + x^r$  et  $g(x) = (a + x)^r$ . Nous avons  $f(0) = g(0) = a^r$ , et, en utilisant la fonction  $f_\alpha$  définie par  $f_\alpha(x) = x^\alpha$ , nous avons

$$f'(x) = rx^{r-1} = rf_{r-1}(x) \quad (27.536a)$$

$$g'(x) = r(a + x)^{r-1} = rf_{r-1}(a + x). \quad (27.536b)$$

Vu que  $r > 1$ , la fonction  $f_{r-1}$  est strictement croissante sur les positifs par la proposition 12.427. Donc pour  $x \geq 0$  nous avons

$$f'(x) = rf_{r-1}(x) \leq rf_{r-1}(a + x) = g'(x). \quad (27.537)$$

Nous avons donc  $f(b) \leq g(b)$  comme souhaité.

Nous passons à la seconde inégalité. Le lemme 17.87 nous dit que la fonction  $f: x \mapsto x^r$  est convexe. Donc

$$f\left(\frac{a}{2} + \frac{b}{2}\right) \leq \frac{1}{2}f(a) + \frac{1}{2}f(b). \quad (27.538)$$

De là nous déduisons

$$\frac{(a + b)^r}{2^r} \leq \frac{1}{2}(a^r + b^r), \quad (27.539)$$

c'est-à-dire la seconde inégalité. □

Comme semble indiquer le code suivant, la seconde inégalité du lemme 27.134 semble vraie pour tous les complexes.

```

1 import random
2
3
4 def f1(z1: complex, z2: complex, r: float) -> float:
5     return abs((z1+z2)/2)**r
6
7
8 def f2(z1: complex, z2: complex, r: float) -> float:
9     return (abs(z1)**r + abs(z2)**r)/2
10
11
12 def test(z1: complex, z2: complex, r: float) -> float:
13     """Should be positive."""
14     return f2(z1, z2, r) - f1(z1, z2, r)
15
16
17 for _ in range(1, 10):
18     r = random.uniform(1, 1.3)
19     x1 = random.uniform(-10, 10)
20     y1 = random.uniform(-10, 10)
21     x2 = random.uniform(-10, 10)
22     y2 = random.uniform(-10, 10)
23     z1 = complex(x1, y1)

```

```

24     z2 = complex(x2, y2)
25     result = test(z1, z2, r)
26     print("-----")
27     print(f"z1={z1}")
28     print(f"z2={z2}")
29     print(f"r={r}")
30     print("----> ", result)
31     if result < 0:
32         print("!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!")

```

tex/frido/codeSnip\_5.py

PROPooPMIGooRAjROX

**Proposition 27.135** ([749]).Soient  $\alpha, \beta \in \mathbb{C}$  et  $r \geq 1$ . Nous avons

$$\left| \frac{\alpha + \beta}{2} \right|^r \leq \frac{|\alpha|^r + |\beta|^r}{2}. \quad (27.540)$$

*Démonstration.* Nous prouvons en montant progressivement en généralité.

- (1)  $\alpha = 1$  et  $\beta = t \geq 1$  SPITEMooUDCCooSxNxhiv Nous commençons avec  $\alpha = 1$  et en supposant que  $\beta$  est un réel plus grand ou égal à 1. Nous devons prouver que

$$\left( \frac{1+t}{2} \right)^r \leq \frac{1+t^r}{2}. \quad (27.541)$$

Nous posons

$$\begin{aligned} f &: [1, \infty[ \rightarrow \mathbb{R} \\ t &\mapsto \frac{1+t^r}{2} - \left( \frac{1+t}{2} \right)^r, \end{aligned} \quad (27.542)$$

et nous prouvons que  $f(t) \geq 0$  pour tout  $t \geq 1$ . Pour cela nous calculons la dérivée :

$$f'(t) = \frac{r}{2} \left( t^{r-1} - \left( \frac{1+t}{2} \right)^{r-1} \right) \geq 0. \quad (27.543)$$

Nous avons utilisé le fait que  $x \mapsto x^{r-1}$  est croissante et que  $t \geq 1$ . Vu que  $f(1) = 1 - 1 = 0$ , la proposition 12.210(2) donne

$$f(t) \geq f(1) = 0 \quad (27.544)$$

pour tout  $t \geq 1$ 

- (2)  $\alpha = z \in \mathbb{C}$  avec  $|z| \geq 1$  et  $\beta = 1$  SPITEMooLSGLooXiltES  
Nous avons

$$\left| \frac{z+1}{2} \right|^r \leq \left( \frac{|z|+1}{2} \right)^r \quad \text{pcq. } |z+1| \leq |z|+1 \quad (27.545a)$$

$$\leq \frac{1+|z|^2}{2} \quad \text{par point (1).} \quad (27.545b)$$

**(3) Cas général**Notons que si  $\beta = 0$  c'est facile. Nous supposons que  $\beta \neq 0$  et que  $|\alpha| \geq |\beta|$ . En posant  $z = \alpha/\beta$  nous avons  $|z| \geq 1$  et nous pouvons calculer un peu :

$$\left| \frac{\alpha + \beta}{2} \right|^r = \left| \frac{\beta \frac{\alpha+\beta}{\beta}}{2} \right|^r \quad (27.546a)$$

$$= |\beta|^r \left| \frac{z+1}{2} \right|^r \quad (27.546b)$$

$$\leq |\beta|^r \frac{1+|z|^2}{2} \quad \text{par le point (2)} \quad (27.546c)$$

$$= \frac{|\beta|^r + |\alpha|^r}{2}. \quad (27.546d)$$

□

PROPooHEWFooJpfzEL

**Proposition 27.136** ([739]).Si  $a, b \geq 0$  et si  $0 < r \leq 1$ , alors

$$(a + b)^r \leq a^r + b^r. \quad (27.547)$$

*Démonstration.* Nous avons un calcul :

$$(a + b)^r - a^r = \int_a^{a+b} rt^{r-1} dt \quad \text{tho. 14.288} \quad (27.548a)$$

$$= \int_0^b r(s + a)^{r-1} ds \quad \text{ch.var. } s = t - a \quad (27.548b)$$

$$\leq \int_0^b rs^{r-1} ds \quad \text{cf. justif. SUBEQooUCTQooBULeqL} \quad (27.548c)$$

$$= b^r. \quad (27.548d)$$

Pour (27.548c), nous savons que  $r < 1$  (le cas  $r = 0$  est facile), de telle sorte que  $r - 1 < 0$ . La fonction  $s \mapsto (s + a)^{r-1}$  est donc décroissante. □

PROPooTIJYooJGxphQ

**Proposition 27.137** (c-r inégalité[739]).

Soit un espace mesuré  $(\Omega, \mathcal{A}, \mu)$ . Soient  $f, g \in L^r(\Omega, \mathcal{A}, \mu)$  des fonctions à valeurs réelles<sup>63</sup>. En posant  $c_r = 1$  pour  $0 < r \leq 1$  et  $c_r = 2^{r-1}$  si  $r \geq 1$ , nous avons

(1) Pour tout  $\omega \in \Omega$ ,

$$|f(\omega) + g(\omega)|^r \leq c_r |f(\omega)|^r + c_r |g(\omega)|^r. \quad \text{EQooWIQOooLWPfoQ} \quad (27.549)$$

(2) Nous avons

$$\int_{\Omega} |f + g|^r d\mu \leq c_r \int_{\Omega} |f|^r d\mu + c_r \int_{\Omega} |g|^r d\mu. \quad \text{EQooAJTKooGNxHat} \quad (27.550)$$

*Démonstration.* L'inégalité (27.549) provient de propositions 27.136 et 27.135 (suivant la valeur de  $r$ ). L'inégalité (27.550) demande maintenant d'intégrer sur  $\Omega$ . □

### 27.11.2 Inégalité de Hanner

Nous allons démontrer les inégalités de Hanner dans le théorème 27.142. Vu que ce sera un peu longuet, nous faisons un lemme.

LEMooDHRCoQispyC

**Lemme 27.138.**Soient  $z_1, z_2 \in \mathbb{C}$ . Nous avons

$$|z_1 + z_2|^p + |z_1 - z_2|^p \geq (|z_1| + |z_2|)^p + ||z_1| - |z_2||^p. \quad \text{EQooMUXVooSpCSyG} \quad (27.551)$$

*Démonstration.* Soient  $z_1, z_2 \in \mathbb{C}$ . Nous posons

$$d = |z_1 + z_2|^p + |z_1 - z_2|^p. \quad \text{EQooJKYZooFzbETG} \quad (27.552)$$

Pour  $|z_1|$  et  $|z_2|$  fixés, nous nous demandons quel est le minimum possible de  $d$ .

Si  $|z_1| = 0$ , alors le minimum est  $2|z_2|^p$  et si  $|z_2| = 0$  alors il est  $2|z_1|^p$ . Pour les autres cas, nous posons  $|z_1| = a > 0$  ainsi que  $b \in \mathbb{R}$  et  $\theta \in \mathbb{R}$  tels que<sup>64</sup>

$$z_2 = z_1 a^{-1} b e^{i\theta}. \quad (27.553)$$

Nous avons déjà que  $z_1 + z_2 = z_1(1 + a^{-1} b e^{i\theta})$  et donc

$$|z_1 + z_2| = a |1 + a^{-1} b e^{i\theta}| = |a + b e^{i\theta}| \quad (27.554)$$

63. Par défaut, les éléments de  $L^p$  sont à valeurs complexe. Et oui, je sais que ce sont des classes de fonctions, et donc elles peuvent être à valeurs complexes sur une partie de mesure nulle.

64. Proposition 18.63

parce que  $a > 0$ . De plus,

$$|a + be^{i\theta}|^2 = (a + be^{i\theta})(a + be^{-i\theta}) = a^2 + b^2 + 2ab \cos(\theta) \quad (27.555)$$

parce que  $e^{i\theta} + e^{-i\theta} = 2\cos(\theta)$ . Nous posons

$$d(\theta) = |a + be^{i\theta}|^p + |a - be^{i\theta}|^p. \quad (27.556)$$

En développant,

$$d(\theta) = (a^2 + b^2 + 2ab \cos(\theta))^{p/2} + (a^2 + b^2 - 2ab \cos(\theta))^{p/2}. \quad (27.557)$$

Trouvons le minimum de cette fonction de  $\theta$ . D'abord sa dérivée :

$$d'(\theta) = pab \sin(\theta) [(a^2 + b^2 - 2ab \cos(\theta))^{p/2-1} - (a^2 + b^2 + 2ab \cos(\theta))^{p/2-1}] \quad (27.558a)$$

$$= pab \sin(\theta) s(\theta). \quad (27.558b)$$

Nous avons  $s(\theta) = 0$  pour  $\theta = \pi/2$  et  $\theta = 3\pi/2$ . Il faut surtout remarquer que  $1 < p < 2$ , ce qui donne  $\frac{p}{2} - 1 < 0$ . La fonction  $x \mapsto x^{p/2-1}$  est donc décroissante. Cela pour dire que

$$s(0) = (|a - b|^2)^{p/2-1} - (|a + b|^2)^{p/2-1} > 0. \quad (27.559)$$

De la même façon,  $s(\pi) = -s(0) < 0$ . Cela permet d'écrire un petit tableau de signe de  $d'$ , et de conclure que  $d(\theta)$  a un minimum en 0 et en  $\pi$ . Calcul fait, nous avons

$$d(0) = d(\pi) = |a + b|^p + |a - b|^p. \quad (27.560)$$

En reliant à (27.552) nous avons l'inégalité

$$|z_1 + z_2|^p + |z_1 - z_2|^p \geq (a + b)^p + |a - b|^p. \quad (27.561) \quad \text{E} \text{QooV} \text{H} \text{Q} \text{QooJ} \text{cheCB}$$

Nous rappelons que  $a = |z_1|$  et que  $z_2 = z_1 a^{-1} b e^{i\theta}$ . Notons au passage que  $|z_2| = b$ , donc que ce que nous dit l'équation (27.561) est que

$$|z_1 + z_2|^p + |z_1 - z_2|^p \geq (|z_1| + |z_2|)^p + ||z_1| - |z_2||^p. \quad (27.562)$$

□

Encore dans la catégorie des lemmes pour les inégalités de Hanner, nous avons celui-ci.

**Lemme 27.139** ([1, 750]).

La fonction

$$\eta: ]0, \infty[ \rightarrow \mathbb{R} \quad (27.563)$$

$$a \mapsto (a^{1/p} + 1)^p + |a^{1/p} - 1|^p$$

est strictement convexe.

*Démonstration.* La fonction  $\eta$  est une fonction de classe  $C^\infty$  sur  $]0, \infty[\setminus\{1\}$ . Quelle est sa régularité en  $a = 1$ ? À cause de la valeur absolue, il n'est pas clair qu'elle y soit dérivable. En tout cas, la fonction  $x \mapsto |x - 1|$  n'est pas dérivable en  $x = 1$ , mais peut-être que les exposants aident à lisser. Nous y reviendrons.

Afin de suivre les calculs nous introduisons quelques fonctions :

$$so(x) = 1 + x^{1/p} \quad (27.564a)$$

$$di(x) = 1 - x^{1/p} \quad (27.564b)$$

$$dj(x) = x^{1/p} - 1 \quad (27.564c)$$

Pour les dérivées, nous avons

$$so'(x) = \frac{1}{p}x^{1/p-1} \quad (27.565a)$$

$$di'(x) = -so'(x) \quad (27.565b)$$

$$dj'(x) = so'(x). \quad (27.565c)$$

Nous divisons les cas selon  $a < 1$  ou  $a > 1$ .

(1) **Pour  $a < 1$**  Nous avons

$$\eta(a) = so(a)^p + di(a)^p, \quad (27.566)$$

et la première dérivée donne :

$$\eta'(a) = p so'(a)(so(a)^{p-1} - di(a)^{p-1}). \quad (27.567)$$

Pour la seconde dérivée nous trouvons d'abord

$$\begin{aligned} \eta''(a) &= \left(\frac{1-p}{p}\right) a^{\frac{1}{p}-2} (so(a)^{p-1} - di(a)^{p-1}) \\ &\quad + \frac{p-1}{p} a^{\frac{2}{p}-2} (so(a)^{p-2} + di(a)^{p-2}). \end{aligned} \quad (27.568)$$

À partir de là, le truc est de substituer les expressions suivantes :

$$so(a)^{p-1} = so(a)^{p-2} so(a) = so(a)^{p-2} + so(a)^{p-2} a^{1/p} \quad (27.569a)$$

$$di(a)^{p-1} = di(a)^{p-2} - a^{1/p} di(a)^{p-2}. \quad (27.569b)$$

Plein de trucs se simplifient et nous obtenons

$$\eta''(a) = \frac{p-1}{p} a^{\frac{1}{p}-2} (di(a)^{p-1} - so(a)^{p-2}). \quad (27.570)$$

(2) **Pour  $a > 1$**  Les calculs sont essentiellement les mêmes, en partant de

$$\eta(a) = so(a)^p + dj(a)^p. \quad (27.571)$$

Les résultats sont :

$$\eta'(a) = p so'(a)(so(a)^{p-1} + dj(a)^{p-1}), \quad (27.572)$$

et

$$\eta''(a) = \frac{p-1}{p} a^{\frac{1}{p}-2} (dj(a)^{p-2} - so(a)^{p-2}). \quad (27.573)$$

(3) **Résumé pour  $a \neq 1$**

Au final, nous avons pour tout  $a \neq 1$  :

$$\eta''(a) = \frac{p-1}{p} a^{\frac{1}{p}-2} (|1 - a^{1/p}|^{p-2} - (1 + a^{1/p})^{p-2}). \quad (27.574)$$

Ce qu'il se passe en  $a = 1$  est encore une question ouverte que nous traitons maintenant.

(4) **Pour  $a = 1$**

Les limites des expressions (27.567) et (27.572) en  $a = 1$  sont vite calculées et c'est  $2^{p-1}$  dans les deux cas. Donc la dérivée admet une prolongation continue en  $a = 1$ . Nous allons prouver que la fonction  $\eta$  est en réalité dérivable en  $a = 1$  et que la dérivée vaut  $2^{p-1}$ .

Nous nous concentrons sur la partie difficile donnée par  $f(x) = |x^{1/p} - 1|^p$ . Elle est donnée par

$$f(x) = \begin{cases} di(x)^p & \text{si } x < 1 \\ dj(x)^p & \text{si } x > 1 \\ 0 & \text{si } x = 1. \end{cases} \quad (27.575)$$

Si  $f'(1)$  existe, alors elle est égale à la limite

$$f'(1) = \lim_{\epsilon \rightarrow 0} \frac{f(1) - f(1 - \epsilon)}{\epsilon}. \quad (27.576)$$

Les deux limites à calculer sont :

$$\lim_{\epsilon \rightarrow 0^+} \frac{((1 + \epsilon)^{1/p} - 1)^p}{\epsilon} \quad (27.577)$$

et

$$\lim_{\epsilon \rightarrow 0^-} \frac{(1 - (1 + \epsilon)^{1/p})^p}{\epsilon}. \quad (27.578)$$

La première se traite par la règle de l'Hospital <sup>65</sup>, et le résultat est zéro. Pour la seconde, il faut juste transformer

$$\lim_{\epsilon \rightarrow 0^+} \frac{((1 + \epsilon)^{1/p} - 1)^p}{\epsilon} = \lim_{h \rightarrow 0^+} \frac{(1 - (1 - h)^{1/p})^p}{-h}, \quad (27.579)$$

qui se traite également par la règle de l'Hospital. Le résultat est également zéro.

Donc  $\eta$  est dérivable en  $a = 1$  et la dérivée vaut  $\eta'(1) = 2^{p-1}$ .

En récapitulant, nous avons  $\eta'' > 0$  sur  $]0, \infty[ \setminus \{1\}$ , donc  $\eta'$  est croissante sur cette partie (proposition 12.209). Vu que  $\eta'$  est continue sur  $]0, \infty[$ , elle est même croissante (strictement) sur tout  $]0, \infty[$ .

La proposition 17.83 conclut que  $\eta$  est strictement convexe sur  $]0, \infty[$ .  $\square$

Toujours dans la catégorie des lemmes pour les inégalités de Hanner, nous avons celui-ci.

**Lemme 27.140** ([750]).

Soit  $1 < p < 2$ . La fonction

$$\begin{aligned} \xi: \mathbb{R}^+ \times \mathbb{R}^+ &\rightarrow \mathbb{R} \\ (a, b) &\mapsto (a^{1/p} + b^{1/p})^p + |a^{1/p} - b^{1/p}|^p \end{aligned} \quad (27.580)$$

est convexe.

Pour rappel, les conventions de données en 1.531 donnent  $\mathbb{R}^+ = [0, \infty[$ .

*Démonstration.* La fonction  $\xi$  vérifie facilement les conditions suivante :

- $\xi(a, b) = \xi(b, a)$ ,
- $\xi(0, 0) = 0$ ,
- $\xi(ta, tb) = t\xi(a, b)$  pour tout  $t \geq 0$ .

Nous posons

$$\begin{aligned} \eta: \mathbb{R}^+ &\rightarrow \mathbb{R} \\ a &\mapsto \xi(a, 1). \end{aligned} \quad (27.581)$$

Le lemme 27.139 dit que  $\eta$  est strictement convexe, et le lemme 17.93 conclut que  $\xi$  est convexe.  $\square$

LEMooWIPYooMZqjbn

**Lemme 27.141** ([751]).

Soit  $1 < p < 2$ . Nous considérons les fonctions

$$\begin{aligned} \alpha: [0, 1] &\rightarrow \mathbb{R} \\ x &\mapsto (1 + x)^{p-1} + (1 - x)^{p-1} \end{aligned} \quad (27.582)$$

et

$$\begin{aligned} \beta: [0, 1] &\rightarrow \mathbb{R} \\ x &\mapsto x^{1-p}((1 + x)^{p-1} - (1 - x)^{p-1}). \end{aligned} \quad (27.583)$$

Soient  $A, B \in \mathbb{R}$ . Nous avons

$$\alpha(x)|A|^p + \beta(x)|B|^p \leq |A + B|^p + |A - B|^p. \quad (27.584)$$

---

65. Proposition 12.219



*Démonstration.* Plusieurs étapes.

- (1)  $\beta(x) \leq \alpha(x)$  Nous avons  $\alpha(1) = \beta(1) = 2^{p-1}$ . Pour les autres valeurs de  $x$ , nous allons raisonner avec la dérivée. La valeur de  $\alpha'(x)$  est facile à calculer

$$\alpha'(x) = (p-1)(x+1)^{p-2} - (p-1)(1-x)^{p-2}. \quad (27.585)$$

Pour  $\beta'(x)$  c'est un peu plus lourd. En substituant  $(1+x)^{p-1} = (1+x)^{p-2}(1+x)$  et  $(1-x)^{p-1} = (1-x)^{p-2}(1-x)$  nous pouvons regrouper les termes en  $(1+x)^{p-2}$  et  $(1-x)^{p-2}$ . Après un peu de travail,

$$\beta'(x) = \frac{p-1}{x^p} ((1-x)^{p-2} - (1+x)^{p-2}). \quad (27.586)$$

Cela nous permet de calculer  $\alpha' - \beta'$  :

$$\alpha'(x) - \beta'(x) = (p-1)\left(1 + \frac{1}{x^p}\right)((1+x)^{p-2} - (1-x)^{p-2}). \quad (27.587)$$

Vu que  $1 < p < 2$ , le nombre  $p-2$  est strictement négatif ; afin de travailler avec des exposants positifs, nous écrivons

$$\alpha'(x) - \beta'(x) = \underbrace{(p-1)}_{>0} \underbrace{\left(1 + \frac{1}{x^p}\right)}_{>0} \underbrace{\left(\frac{1}{(1+x)^{2-p}} - \frac{1}{(1-x)^{2-p}}\right)}_{<0}. \quad (27.588)$$

Nous avons  $\alpha'(x) - \beta'(x) < 0$  pour tout  $x \in ]0, 1]$ . Du fait qu'en plus nous avons  $\alpha(1) = \beta(1)$ , nous déduisons que  $\alpha(x) \geq \beta(x)$ .

- (2) **Une petite étude de fonction** Soit  $R \in [0, 1]$ . Nous considérons la fonction

$$\begin{aligned} F: [0, 1] &\rightarrow \mathbb{R} \\ x &\mapsto \alpha(x) + R^p \beta(x). \end{aligned} \quad (27.589)$$

Nous montrons maintenant que cette fonction a un maximum global pour  $x = R$ . D'abord sa dérivée :

$$F'(x) = \underbrace{(p-1)}_{>0} \underbrace{\left((1-x)^{p-1} - (1+x)^{p-2}\right)}_{<0} \left(1 - \left(\frac{R}{x}\right)^p\right) \quad (27.590)$$

Nous avons

- $F'(x) = 0$  pour  $x = R$ ,
- $F'(x) < 0$  pour  $x > R$ ,
- $F'(x) > 0$  pour  $x < R$ .

Donc  $x = R$  est bien un maximum global.

- (3) **Pause** Nous avons les petits résultats utiles pour commencer à prouver. Petite pause avant de commencer ; pas de panique, ça ne va pas être trop violent.
- (4) **Pour  $0 < B < A$**  Nous devons prouver que

$$\alpha(x)A^p + \beta(x)B^p \leq (A+B)^p + (A-B)^p. \quad (27.591) \quad \text{EQooEPKRooBYJDSE}$$

En divisant par  $A^p$  et en posant  $R = B/A$ , l'inéquation (27.591) est équivalente à

$$\alpha(x) + \beta(x)R^p \leq (1+R)^p + (1-R)^p \quad (27.592)$$

où  $R \in ]0, 1[$  parce que nous avons supposé  $0 < B < A$ . Nous avons (il y a un petit calcul pour  $F(R)$ )

$$(1+R)^p + (1-R)^p = F(R) \geq F(x) = \alpha(x) + \beta(x)R^p. \quad (27.593)$$

ok.

(5) Pour  $0 < A < B$  Lorsque  $0 < A < B$  nous avons

$$\begin{aligned}\alpha(x)|A|^p + \beta(x)|B|^p &= \alpha(x)A^p + \beta(x)B^p & (27.594a) \\ &\leq \alpha(x)B^p + \beta(x)A^p & (27.594b) \\ &\leq (B+A)^p + (B-A)^p & (27.594c) \\ &= |A+B|^p + |A-B|^p. & (27.594d)\end{aligned}$$

Justification :

- Pour (27.594b), c'est parce que  $\alpha(x) > \beta(x)$  ; alors en mettant le plus grand de  $A$  et  $B$  devant le  $\alpha$  au lieu du  $\beta$ , nous majorons.
  - Pour (27.594c), c'est l'inégalité dans le cas  $0 < B < A$ , mais en inversant les noms de  $A$  et  $B$ .
- (6) Pour  $0 < A = B$  Toutes les expressions sont continues par rapport à  $B$  (fixons  $x$  et  $A$ ). Nous avons prouvé pour  $B < A$  et pour  $B > A$ . Par continuité, l'inégalité est encore valide pour  $A = B$ .
- (7) Pour  $A < 0, B > 0$  En posant  $A' = -A$  nous avons  $A' > 0$  et nous pouvons écrire

$$|A+B|^p + |A-B|^p = |-A'+B|^p + |-A'-B|^p = |B-A'|^p + |B+A'|^p \geq \alpha(x)|A'|^p + \beta(x)|B|^p. \quad (27.595)$$

Nous avons utilisé, avec  $A'$  et  $B$  le cas déjà prouvé  $A', B > 0$ .

- (8) Pour  $A > 0, B < 0$  Celui-là, je vous le laisse.
- (9) Pour  $A < 0, B < 0$  Posez  $A' = -A$  et  $B' = -B$  et hop.

□

THOooZRRYooBTBQKW

**Théorème 27.142** (Inégalités de Hanner[750, 751]).

Soit un espace mesuré  $(\Omega, \mathcal{A}, \mu)$ . Soit  $1 < p < 2$  et  $f, g \in L^p(\Omega, \mathcal{A}, \mu)$  ; nous avons

$$(\|f\|_p + \|g\|_p)^p + \left| \|f\|_p - \|g\|_p \right|^p \leq \|f+g\|_p^p + \|f-g\|_p^p \quad (27.596)$$

Il y a égalité si et seulement si  $f(t)$  et  $g(t)$  sont colinéaires pour presque tout  $t$ .

*Démonstration.* Nous supposons que  $\|f\|_p \geq \|g\|_p$  pour fixer les idées. De toutes façons, la symétrie des formules nous fait passer de ce cas à l'autre sans difficulté.

Soit  $x \in [0, 1]$ . Nous écrivons l'inégalité du lemme 27.141 pour  $A = |f(\omega)|$  et  $B = |g(\omega)|$  :

$$\alpha(x)|f(\omega)|^p + \beta(x)|g(\omega)|^p \leq |f(\omega) + g(\omega)|^p + |f(\omega) - g(\omega)|^p. \quad (27.597)$$

Nous intégrons cela par rapport à  $\omega$  sur  $\Omega$  :

$$\alpha(x)\|f\|_p^p + \beta(x)\|g\|_p^p \leq \|f+g\|_p^p + \|f-g\|_p^p. \quad (27.598)$$

Et là vient l'idée qu'on se demande ce qui est passé par l'esprit du mec qui a tout combiné : nous évaluons cela pour  $x = \frac{\|g\|_p^p}{\|f\|_p^p}$ , ce qui est permis parce que nous avons supposé  $\|f\|_p \geq \|g\|_p$ . Faites le calcul, collectez les termes identiques, vous obtiendrez

$$(\|f\|_p + \|g\|_p)^p + (\|f\|_p - \|g\|_p)^p \leq \|f+g\|_p^p + \|f-g\|_p^p. \quad (27.599)$$

Et vu que  $\|f\|_p \geq \|g\|_p$ , nous pouvons gratuitement faire

$$\|f\|_p - \|g\|_p = \left| \|f\|_p - \|g\|_p \right|. \quad (27.600)$$

Fin pour Hanner.

□

### 27.11.3 Inégalités de Clarkson

**Lemme 27.143** ([752]).

Si  $p \geq 2$  et si  $a, b \in \mathbb{C}$ , alors

$$\left| \frac{a+b}{2} \right|^p + \left| \frac{a-b}{2} \right|^p \leq \frac{1}{2}(|a|^p + |b|^p). \quad (27.601)$$

*Démonstration.* Nous prouvons l'inégalité en montant petit à petit en généralité.

(1) Avec  $x > 0$  Soit  $x \geq 0$ . Nous montrons dans cette partie l'inégalité

$$x^p + 1 \leq (x+1)^{p/2}. \quad (27.602)$$

Pour cela nous considérons la fonction

$$\begin{aligned} f: [0, \infty[ &\rightarrow \mathbb{R} \\ t &\mapsto (t^2 + 1)^{p/2} - t^p - 1. \end{aligned} \quad (27.603)$$

Nous avons  $f(0) = 0$ , mais aussi, en utilisant les règles de dérivation<sup>66</sup> nous trouvons vite

$$f'(t) = p(t^2 + 1)^{p/2-1}t - pt^{p-1}. \quad (27.604)$$

Vu que  $(t^2 + 1)^{p/2-1} \geq t^{p-2}$ , le signe de  $f'(t)$  est toujours strictement positif pour  $t > 0$ . La proposition 12.209 fait que  $f$  est strictement croissante et que  $f(t) > 0$  pour tout  $t > 0$ .

(2) Avec  $x, y \geq 0$  Soient  $x, y \geq 0$  dans  $\mathbb{R}$ . Nous prouvons dans cette partie que

$$(x^2 + y^2)^{p/2} \geq x^p + y^p. \quad (27.605)$$

Il s'agit d'appliquer l'inégalité (27.602) à  $x/y$  :

$$\left( \left( \frac{x}{y} \right)^2 + 1 \right)^{p/2} \geq \left( \frac{x}{y} \right)^p + 1. \quad (27.606)$$

En multipliant par  $y^p$  et en simplifiant un peu, nous trouvons le résultat (27.605).

(3) Avec  $a, b \in \mathbb{C}$  Nous appliquons l'inégalité (27.605) à  $x = |\frac{a+b}{2}|$  et  $y = |\frac{a-b}{2}|$ . Cela donne :

$$\left| \frac{a+b}{2} \right|^p + \left| \frac{a-b}{2} \right|^p \leq \left( \left| \frac{a+b}{2} \right|^2 + \left| \frac{a-b}{2} \right|^2 \right)^{p/2} \quad (27.607a)$$

$$= \left( \frac{2|a|^2 + 2|b|^2}{4} \right)^{p/2} \quad (27.607b)$$

$$\leq \frac{1}{2}|a|^p + \frac{1}{2}|b|^p. \quad (27.607c)$$

La dernière ligne est la convexité de la fonction  $t \mapsto t^{p/2}$  (lemme 17.87).

□

**Lemme 27.144.**

Si  $1 < p < 2$ , alors l'exposant conjugué  $q$  vérifie  $q > 2$ .

*Démonstration.* Nous considérons  $q$  en fonction de  $p$ , sur le domaine  $1 < p < 2$  :

$$q(p) = \frac{p}{p-1}. \quad (27.608)$$

66. Par exemple celle de la proposition 14.295.

Donc <sup>67</sup>  $q(1) = \infty$  et  $q(2) = 2$ . Nous étudions ensuite la dérivée :

$$q'(p) = -\frac{1}{(p-1)^2} < 0. \quad (27.609)$$

C'est donc une fonction strictement décroissante. Vues les valeurs aux bornes, nous voyons que  $q(p) > 2$  sur tout son domaine.  $\square$

LEMooMKIXooVOYaxI

**Lemme 27.145** ([753]).

Soit  $1 < p < 2$ . Pour  $x, y \in \mathbb{R}$  nous avons

$$|x + y|^q + |x - y|^q \leq 2(|x|^p + |y|^p)^{q-1}. \quad (27.610)$$

*Démonstration.* Nous considérons l'exposant conjugué  $q$  et  $p$ , c'est-à-dire  $q$  tel que  $\frac{1}{p} + \frac{1}{q} = 1$ . Nous considérons la fonction

$$\begin{aligned} f: \mathbb{R} \times [0, 1] &\rightarrow \mathbb{R} \\ (\alpha, z) &\mapsto (1 + \alpha^{1-q}z)(1 + \alpha z)^{q-1} + (1 - \alpha^{1-q}z)(1 - \alpha z)^{q-1}. \end{aligned} \quad (27.611)$$

Cette fonction vérifie

$$f(1, z) = (1 + z)^q + (1 - z)^q, \quad (27.612)$$

EQooRFZQooJydocT

ainsi que <sup>EQooISBRooHMiPRE</sup>

$$f(z^{p-1}, z) = (1 + z^{(p-1)(1-q)})(1 + z^p)^{q-1} + (1 - z^{(p-1)(1-q)})(1 + z^p)^{q-1} \quad (27.613a)$$

$$= 2(1 + z^p)^{q-1}. \quad (27.613b)$$

Nous montrons maintenant que  $(\partial_\alpha f)(\alpha, z) \leq 0$  pour tout  $\alpha \in ]0, 1[$  et pour tout  $z \in ]0, 1[$ . C'est du calcul :

$$\frac{\partial f}{\partial \alpha}(\alpha, z) = (1-q)z[\alpha^{-q}(1+\alpha z)^{q-1} - (1+\alpha^{1-q}z)(1+\alpha z)^{q-1} - \alpha^{-q}(1-\alpha z)^{q-1} + (1-\alpha^{1-q}z)(1-\alpha z)^{q-2}]. \quad (27.614)$$

Maintenant nous factorisons  $(1+\alpha z)^{q-2}$  grâce à la décomposition  $(1+\alpha z)^{q-1} = (1+\alpha z)(1+\alpha z)^{q-2}$ . Notez que le lemme 27.144 donne  $q > 2$ , et donc pas de problèmes avec la puissance  $q-2$ . Nous continuons le calcul

$$\frac{\partial f}{\partial \alpha}(\alpha, z) = (1-q)z(1+\alpha z)^{q-2}[\alpha^{-q}(1+\alpha z) - (1+\alpha^{1-q}z)] \quad (27.615a)$$

$$\begin{aligned} &+ (1-q)z(1-\alpha z)^{q-2}[-\alpha^{-q}(1-\alpha z) + (1-\alpha^{1-q}z)] \\ &= (1-q)z(1+\alpha z)^{q-2}[\alpha^{-q} + \alpha^{-q+1}z - 1 - \alpha^{1-q}z] \end{aligned} \quad (27.615b)$$

$$\begin{aligned} &+ (1-q)z(1-\alpha z)^{q-2}[-\alpha^{-q} + \alpha^{-q+1} + 1 - \alpha^{1-q}z] \\ &= (1-q)z(1+\alpha z)^{q-2}[\alpha^{-q} - 1] + (1-q)z(1-\alpha z)^{q-2}[1 - \alpha^{-q}] \end{aligned} \quad (27.615c)$$

$$= (1-q)z(\alpha^{-q} - 1)[(1+\alpha z)^{q-2} - (1-\alpha z)^{q-2}]. \quad (27.615d)$$

Vu que  $q > 2$ , la fonction  $x \mapsto x^{q-2}$  est strictement croissante sur les positifs <sup>68</sup>. Et vu que  $\alpha z < 1$ , le crochet est strictement positif. Par ailleurs,  $z > 0$ ,  $(1-q) < 0$  et  $(\alpha^{-q} - 1) > 0$  donc nous avons prouvé que  $(\partial_\alpha f)(\alpha, z) \leq 0$ .

Donc  $f$  est décroissante par rapport à  $\alpha$ . Vu que  $z \in [0, 1]$  et que  $p > 1$ , nous avons  $z^{p-1} < 1$  et donc  $f(1, z) \leq f(z^{p-1}, z)$ . Nous y substituons les valeurs calculées en (27.612) et (27.613) :

$$(1 + z)^q + (1 - z)^q \leq 2(1 + z^p)^{q-1}. \quad (27.616)$$

EQooFQJAooPCYtMG

Nous pouvons maintenant facilement prouver notre inégalité.

67. Dire que  $q(1) = \infty$  est un abus de notations pour parler de la limite  $p \rightarrow 1$  avec  $p > 1$ .

68. Proposition 12.427.

- (1) **Pour**  $0 < x < y$  Si  $0 < x < y$ , nous avons  $x/y \in [0, 1]$  et nous pouvons appliquer l'inégalité (27.616) à  $z = x/y$ . Nous avons successivement :

$$\left(1 + \frac{x}{y}\right)^q + \left(1 - \frac{x}{y}\right)^q \leq \left(1 + \frac{x^p}{y^p}\right)^{q-1} \quad (27.617a)$$

$$\left(\frac{x+y}{y}\right)^q + \left(\frac{y-x}{y}\right)^q \leq 2 \left(\frac{y^p + x^p}{y^p}\right)^{q-1} \quad (27.617b)$$

$$y^{-q}(x+y)^q + y^{-q}(y-x)^q \leq 2y^{-p(q-1)}(y^p + x^p) \quad (27.617c)$$

$$(x+y)^q + (y-x)^q \leq 2y^{-p(q-1)+q}(y^p + x^p). \quad (27.617d)$$

Nous avons utilisé la proposition 12.432 sur la composition de puissances. Maintenant il suffit de remarquer que  $q = p(q-1)$  pour avoir le résultat.

- (2)  $x < 0$  et  $y > 0$  En posant  $x' = -x$  nous avons

$$|x+y|^q + |x-y|^q = |-x'+y|^q + |-x'-y|^q \quad (27.618a)$$

$$= |x'-y|^q + |x'+y|^q \quad (27.618b)$$

$$\leq 2(|x'|^p + |y|^p) \quad (27.618c)$$

$$\leq 2(|x|^p + |y|^p). \quad (27.618d)$$

- (3) **Les autres cas** Je vous prie de faire la liste, et d'adapter.

□

Pour d'autres preuves du lemme suivant, voir [754].

LEMooLTR0ooVusGte

**Lemme 27.146** ([755]).

Soient  $a, b \in \mathbb{C}$  ainsi que  $1 < p < 2$ . Nous notons  $q$  l'exposant conjugué de  $p$ . Nous avons l'inégalité

$$|a+b|^q + |a-b|^q \leq 2(|a|^p + |b|^p)^{q-1}. \quad (27.619)$$

*Démonstration.* Commençons doucement avec le cas  $b = 0$ . À gauche nous gardons  $2|a|^q$ , et pour le membre de droite nous remarquons que

$$q-1 = \frac{1}{p-1}, \quad (27.620)$$

de telle sorte que

$$(|a|^p)^{q-1} = |a|^{p/(p-1)} = |a|^q. \quad (27.621)$$

Gardez en tête que, par le lemme 27.144,  $q > 2$ ; ce sera utile.

Nous commençons le vrai combat. Vu que  $|a-b| = |b-a|$  nous pouvons supposer  $|a| \geq |b|$  pour fixer les idées. En utilisant le lemme 18.64, il existe  $t_0 \in [0, \frac{\pi}{2}]$  tel que

$$\begin{aligned} |a+b|^2 &= |a|^2 + |b|^2 + 2|a||b|\cos(t_0) \\ |a-b|^2 &= |a|^2 + |b|^2 - 2|a||b|\cos(t_0) \end{aligned} \quad (27.622)$$

Nous considérons la fonction

$$\begin{aligned} f: [0, 2\pi] &\rightarrow \mathbb{R} \\ t &\mapsto \left(|a|^2 + |b|^2 + 2|a||b|\cos(t)\right)^{q/2} + \left(|a|^2 + |b|^2 - 2|a||b|\cos(t)\right)^{q/2}. \end{aligned} \quad (27.623)$$

Elle vérifie

$$f(t_0) = (|a+b|^2)^{q/2} + (|a-b|^2)^{q/2}. \quad (27.624)$$

Vu que  $|a+b|$  et  $|a-b|$  sont positifs, nous pouvons « simplifier » le carré et la racine carré, de telle sorte que

$$f(t_0) = |a+b|^q + |a-b|^q. \quad (27.625)$$

Nous cherchons un maximum pour  $f$  sur  $[0, \frac{\pi}{2}]$ . Pour cela, nous prenons d'abord la dérivée :

$$f'(t) = -q|a||b|\sin(t) \left[ (|a|^2 + |b|^2 + 2|a||b|\cos(t))^{q/2-1} - (|a|^2 + |b|^2 - 2|a||b|\cos(t))^{q/2-1} \right]. \quad (27.626)$$

Notez que  $q > 2$ , donc la fonction  $x \mapsto x^{q/2-1}$  est croissante.

Pour  $t \in [0, \pi/2]$ , nous avons  $\cos(t) \geq -\cos(t)$  ainsi que  $\sin(t) \geq 0$ . Donc  $f'(t) \leq 0$ . De la même manière, nous avons  $f'(t) \geq 0$  pour  $t \in [\pi/2, \pi]$ .

Par le lien entre dérivée et croissance (proposition 12.210), nous savons que le maximum de  $f$  sur  $[0, \pi]$  est atteint en 0 ou en  $\pi$ .

Nous avons, en utilisant la supposition  $|a| \geq |b|$  :

$$f(0) = f(\pi) = (|a|^2 + |b|^2 + 2|a||b|)^{q/2} + (|a|^2 + |b|^2 - 2|a||b|)^{q/2} \quad (27.627a)$$

$$= (|a| + |b|)^q + (|a| - |b|)^q \quad (27.627b)$$

$$= |a| + |b|^q + |a| - |b|^q. \quad (27.627c)$$

En particulier  $f(t_0) \leq f(0)$  et donc

$$|a + b|^q + |a - b|^q \leq |a| + |b|^q + |a| - |b|^q \leq 2(|a|^p + |b|^p)^{q-1}. \quad (27.628)$$

La dernière inégalité est le lemme 27.145 appliqué aux réels  $|a|$  et  $|b|$ . □

PROPOJDOQooWsGlr

**Proposition 27.147** (Inégalité de Clarkson[753]).

Soient  $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ .

(1) Si  $p \geq 2$ , alors

$$\left\| \frac{f+g}{2} \right\|_p^p + \left\| \frac{f-g}{2} \right\|_p^p \leq \frac{1}{2} (\|f\|_p^p + \|g\|_p^p). \quad (27.629)$$

EQooBWDJooCXzdxz

(2) Si  $1 < p < 2$  et si  $q$  est l'exposant conjugué de  $p$ , alors

$$\|f+g\|_p^q + \|f-g\|_p^q \leq 2 (\|f\|_p^p + \|g\|_p^p)^{q-1}, \quad (27.630)$$

EQooXMWBBooYrvaoV

ou

$$\left\| \frac{f+g}{2} \right\|_p^q + \left\| \frac{f-g}{2} \right\|_p^q \leq 2^{1-q} (\|f\|_p^p + \|g\|_p^p)^{q-1}. \quad (27.631)$$

EQooZCWDooBnaMoo

*Démonstration.* En deux parties.

(1) **Pour  $p \geq 2$**  Soient  $f, g \in L^p(\Omega, \mathcal{A}, \mu)$  ; ce sont des fonctions à valeurs dans  $\mathbb{C}$ . Pour chaque  $\omega \in \Omega$  nous considérons les nombres complexes  $f(\omega)$  et  $g(\omega)$  ; nous pouvons écrire l'inégalité du lemme 27.143 :

$$\left| \frac{f(\omega) + g(\omega)}{2} \right|^p + \left| \frac{f(\omega) - g(\omega)}{2} \right|^p \leq \frac{1}{2} (|f(\omega)|^p + |g(\omega)|^p). \quad (27.632)$$

EQooIGNKooKFUpKO

Nous avons les substitutions évidentes  $f(\omega) + g(\omega) = (f+g)(\omega)$  et  $f(\omega) - g(\omega) = (f-g)(\omega)$ . En intégrant alors (27.632) sur  $\Omega$  nous trouvons l'inégalité demandée.

(2) **Pour  $1 < p < 2$**  Il s'agit de faire la même chose, en utilisant l'inégalité de Clarkson du lemme 27.146.

Pour obtenir (27.631), il s'agit simplement de multiplier et diviser le membre de gauche de (27.630) par  $2^q$ .

□

## 27.12 Théorème de la projection normale

### 27.12.1 Espace uniformément convexe

DEFooOPQBooBhufew

**Définition 27.148** (Espace uniformément convexe[756]).

Un espace de Banach  $B$  est **uniformément convexe** si il existe une fonction  $\delta: ]0, \infty[ \rightarrow \mathbb{R}^+$  telle que si

$$(1) \|x\| \leq \|y\| \leq 1,$$

$$(2) \|x - y\| \geq \epsilon,$$

alors

$$\left\| \frac{x + y}{2} \right\| \leq \|y\| - \delta(\epsilon). \quad (27.633)$$

**Lemme 27.149** ([1]).

Si  $B$  est un espace de Banach uniformément convexe, alors pour tout  $k > 0$ , il existe une fonction  $\delta_k: ]0, \infty[ \rightarrow \mathbb{R}^2$  telle que si

$$(1) \|x\| \leq \|y\| \leq k,$$

$$(2) \|x - y\| \geq \epsilon,$$

alors

$$\left\| \frac{x + y}{2} \right\| \leq \|y\| - \delta_k(\epsilon). \quad (27.634)$$

*Démonstration.* Nous posons  $x' = x/k$  et  $y' = y/k$ . Nous avons alors

$$\|x' - y'\| = \frac{\|x - y\|}{k} > \frac{\epsilon}{k}. \quad (27.635)$$

L'uniforme convexité de  $B$  dit alors que

$$\left\| \frac{x' + y'}{2} \right\| > \|y'\| - \delta(\epsilon/k). \quad (27.636)$$

En multipliant cette inégalité par  $k$  nous trouvons

$$\left\| \frac{x + y}{2} \right\| > \|y\| - k\delta(\epsilon/k). \quad (27.637)$$

Donc en posant  $\delta_k(\epsilon) = k\delta(\epsilon/k)$ , nous avons le résultat escompté.  $\square$

DEFooMYYLooJyACPL

**Définition 27.150** (Projection normale[756]).

Soient un espace de Banach  $B$  ainsi que  $V \subset B$ . Soit  $a \in B$ . La fonction

$$\begin{aligned} f: V &\rightarrow \mathbb{R} \\ x &\mapsto d(x, a) \end{aligned} \quad (27.638)$$

possède un infimum<sup>69</sup>  $m$ . Si  $x \in V$  est tel que  $d(x, a) = m$ , alors  $x$  est une **projection normale** de  $a$  sur  $V$ .

PROFoooDKXVoooUoYPgz

**Proposition 27.151** ([756]).

Soient un espace de Banach  $B$  et un sous-espace vectoriel  $V \subset B$ . Si une projection normale de  $a \in B$  sur  $V$  existe, alors elle est unique.

*Démonstration.* Soient deux projections normales  $b, b'$  de  $a$  sur  $V$ .

Si  $m = 0$ , alors  $\|a - b\| = 0$  et  $\|a - b'\| = 0$ , ce qui donne  $a = b$  et  $a = b'$ . Donc d'accord pour  $b = b'$ .

Si  $m > 0$  alors nous utilisons l'inégalité  $\|x + y\| \leq \|x\| + \|y\|$  sous la forme

$$\left\| a - \frac{b + b'}{2} \right\| = \left\| \frac{a - b}{2} + \frac{a - b'}{2} \right\| \leq \left\| \frac{a - b}{2} \right\| + \left\| \frac{a - b'}{2} \right\| = \frac{m}{2} + \frac{m}{2} = m. \quad (27.639)$$

EQoooQWJWoooVaWMCL

69. Toute fonction à valeurs positives possède un infimum, c'est la proposition 1.567.

Mais  $\frac{b+b'}{2} \in V$ , donc

$$\left\| a - \frac{b+b'}{2} \right\| \geq m. \quad (27.640)$$

Nous en déduisons que dans (27.639), toutes les inégalités sont des égalités et en particulier

$$\left\| \frac{b+b'}{2} - a \right\| = m. \quad (27.641)$$

Nous avons donc les deux égalités suivantes :

$$2m = \|a - b\| + \|a - b'\| \quad (27.642)$$

et

$$2m = \|b + b' - 2a\|. \quad (27.643)$$

Cela donne

$$\|a - b\| + \|a - b'\| = \|(a - b) + (a - b')\|. \quad (27.644)$$

Vu que  $B$  est strictement convexe, cela n'est possible que si  $a - b = a - b'$ , ce qui signifie que  $b = b'$ .  $\square$

THOooOOVVooMhzHqd

**Théorème 27.152** ([753, 756]).

Si  $B$  est un espace de Banach uniformément convexe, si  $V \subset B$  est un sous-espace vectoriel fermé et si  $a \in B$ , alors  $a$  admet une unique projection normale<sup>70</sup> sur  $V$ .

*Démonstration.* En deux parties.

- (1) **Unicité** Soient deux projections normales  $b$  et  $b'$  de  $a$  sur  $V$ . Nous avons  $\|a - b\| = \|a - b'\| = m$ . Si  $b \neq b'$ , il existe  $\epsilon > 0$  tel que

$$\|b - b'\| > \epsilon > 0. \quad (27.645)$$

En posant

$$x = \frac{1}{2} \frac{b - a}{m}, \quad y = \frac{1}{2} \frac{b' - a}{m}, \quad (27.646)$$

nous avons  $\|x\| = \|y\| = \frac{1}{2} < 1$ . L'uniforme convexité de  $B$  donne alors

$$\left\| \frac{x+y}{2} \right\| \leq \|y\| - \delta(\epsilon). \quad (27.647)$$

Mais

$$x + y = \frac{\frac{1}{2}(b + b') - a}{m} \quad (27.648)$$

et

$$\|x + y\| \leq 2\|y\| - 2\delta(\epsilon) = 1 - 2\delta(\epsilon) < 1. \quad (27.649)$$

Nous avons donc prouvé que

$$\left\| \frac{1}{2}(b + b') - a \right\| < m, \quad (27.650)$$

ce qui est impossible parce que cela dirait que  $\frac{b+b'}{2}$  est une « meilleure » projection normale que  $b$  et  $b'$ .

- (2) **Existence** Soient  $b_k$  dans  $V$  tels que  $\|a - b_k\| \rightarrow m$ . Nous supposons (quitte à passer à une sous-suite) que

$$\|a - b_{k+1}\| \leq \|a - b_k\|. \quad (27.651)$$

---

70. Définition 27.150.



(2a) **La suite  $(b_k)$  converge**

Nous supposons qu'elle ne converge pas. Elle n'est donc pas de Cauchy parce que  $B$  est de Banach<sup>71</sup> et donc complet. Il existe  $\epsilon > 0$  tel que pour tout  $N \in \mathbb{N}$  il existe  $p, q > N$  tels que

$$\|b_p - b_q\| > \epsilon. \quad (27.652)$$

Nous effectuons quelque choix.

(2a.i) nous choisissons  $q > p$  de telle sorte que  $\|a - b_p\| \leq \|a - b_q\|$ ,

(2a.ii) nous choisissons  $N$  assez grand pour avoir

$$\|a - b_p\| \leq \|a - b_q\| < m. \quad (27.653)$$

Nous avons  $\|(a - b_p) - (a - b_q)\| = \|b_q - b_p\| > \epsilon$ , ce qui avec l'uniforme convexité donne

$$\frac{\|(a - b_p) + (a - b_q)\|}{2} \leq \|a - b_q\| - \delta(\epsilon). \quad (27.654)$$

Donc

$$m \leq \|a - \frac{b_p - b_q}{2}\| = \left\| \frac{(a - b_p) + (a - b_q)}{2} \right\| \leq \|a - b_q\| - \delta(\epsilon) < m - \delta(\epsilon) < m. \quad (27.655)$$

Cela signifie que  $m < m$ , ce qui est impossible.

(2b) **Conclusion** La suite  $(b_k)$  converge dans  $B$ . Vu que  $V$  est fermé, la limite est dans  $V$ . Cette limite, que nous nommons  $b$ , vérifie

$$\|a - b\| \leq \|a - b_k\| \quad (27.656)$$

pour tout  $k$ . Mais comme nous avons  $m \leq \|a - b_k\| \rightarrow m$ , nous avons  $\|a - b\| = m$ , c'est-à-dire que  $b$  est une projection normale de  $a$  sur  $V$ .

□

### 27.12.2 Uniforme convexité des espaces de Lebesgue

PROPooFNLJooDlyIKV

**Proposition 27.153** ([757]).

Si  $1 < p < \infty$ , l'espace  $L^p(\Omega, \mathcal{A}, \mu)$  est uniformément convexe<sup>72</sup>.

*Démonstration.* En deux parties.

(1)  $1 < p \leq 2$  Nous montrons que la fonction  $\delta(\epsilon) = 2^{-q}\epsilon^q$  fonctionne.

Soient  $f, g \in L^p$  telles que  $\|f\|_p \leq \|g\|_p \leq 1$  et  $\|f - g\|_p \geq \epsilon$ . Nous commençons par écrire l'inégalité de Clarkson (27.630) :

$$\left\| \frac{f+g}{2} \right\|_p^q + \left\| \frac{f-g}{2} \right\|_p^q \leq 2^{1-q} (\|f\|_p^p + \|g\|_p^p)^{q-1}. \quad (27.657) \quad \text{EQooOWVEooCGfCuy}$$

Par hypothèse,  $\|f\|_p$  et  $\|g\|_p$  sont plus petites que 1. Vu que  $p > 1$ , nous avons

$$\|f\|_p^p + \|g\|_p^p \leq 1 + 1 = 2. \quad (27.658)$$

En remplaçant dans le membre de droite de (27.657) nous avons

$$\left\| \frac{f+g}{2} \right\|_p^q + \left\| \frac{f-g}{2} \right\|_p^q \leq 2^{1-q} 2^{q-1} = 1, \quad (27.659)$$

et donc

$$\left\| \frac{f+g}{2} \right\|_p^q \leq 1 - \left\| \frac{f-g}{2} \right\|_p^q. \quad (27.660) \quad \text{EQooKARVooDrQuIJ}$$

71. Définition 7.315.

72. Définition 27.148.

Par ailleurs nous avons supposé  $\|f - g\|_p \geq \epsilon$ . Donc aussi <sup>73</sup>

$$\left\| \frac{f - g}{2} \right\|_p^q \geq 2^{-q} \epsilon^q. \quad \text{EQooCGDDooWtDokf} \quad (27.661)$$

Et par un autre ailleurs,

$$\left\| \frac{f + g}{2} \right\|_p = \frac{1}{2} \|f + g\|_p \leq \frac{1}{2} (\|f\|_p + \|g\|_p) \leq 1. \quad \text{EQooOFWYooLVrNDc} \quad (27.662)$$

Vu que nous avons  $q \geq 2$ , cela donne aussi

$$\left\| \frac{f + g}{2} \right\|_p \leq \left\| \frac{f + g}{2} \right\|_p^q. \quad \text{EQooGMPRooGILSss} \quad (27.663)$$

Avec les inégalités (27.661) et 27.663 nous finissons l'inégalité (27.660) :

$$\left\| \frac{f + g}{2} \right\|_p \leq \left\| \frac{f + g}{2} \right\|_p^q \leq 1 - 2^{-q} \epsilon^q \leq \|g\|_p - \delta(\epsilon). \quad (27.664)$$

Okay, c'est bon.

- (2)  $2 \leq p < \infty$  Il s'agit de faire la même chose en partant de Clarkson (27.629). Le résultat est que la fonction  $\delta(\epsilon) = (\epsilon/2)^p$ , ça fonctionne.

□

### 27.12.3 Théorème de la projection normale

#### Proposition 27.154.

Si  $1 < p < \infty$ , et si  $V$  est un sous-espace vectoriel fermé de  $L^p(\Omega, \mathcal{A}, \mu)$ , alors la projection normale <sup>74</sup> de  $a \in L^p$  sur  $V$  existe et est unique.

*Démonstration.* La proposition 27.153 nous indique que l'espace  $L^p(\Omega, \mathcal{A}, \mu)$  est uniformément convexe. Or le théorème 27.152 nous indique que les espaces uniformément convexes vérifient la présente proposition. □

Nous pouvons donner une preuve directe, sans passer par l'uniforme convexité, dans les cas  $p \geq 2$ .

#### Théorème 27.155 (Théorème de la projection normale[758]).

Nous considérons  $p \geq 2$ . Soit un sous-espace vectoriel fermé  $W \subset L^p(\Omega, \mathcal{A}, \mu)$  et  $u_0 \in L^p$ . Nous notons

$$d(u_0, W) = \inf_{w \in W} d(u_0, w). \quad (27.665)$$

Alors il existe  $w_0 \in W$  tel que  $\|u_0 - w_0\| = d(u_0, W)$ .

*Démonstration.* Nous allons séparer trois cas :  $p = 2$  et  $p > 2$ .

- (1)  $p = 2$  Pour  $p = 2$ , nous savons que  $L^2$  est un espace de Hilbert <sup>75</sup>, et nous avons déjà le théorème de la projection 25.7.
- (2)  $p > 2$  Pour chaque  $x \in \Omega$  nous avons  $f(x), g(x) \in \mathbb{C}$  et donc l'identité du parallélogramme <sup>76</sup> :

$$|f(x) - g(x)|^2 + |f(x) + g(x)|^2 = 2|f(x)|^2 + 2|g(x)|^2. \quad \text{EQooUBFEooDUjLub} \quad (27.666)$$

73. Ici j'ai un coefficient un peu différent que celui de [757]. Écrivez-moi pour confirmer ou infirmer mes calculs.

74. Définition 27.150.

75. Lemme 27.105.

76. Théorème 10.1 en remarquant que  $(z_1, z_2) \mapsto z_1 \bar{z}_2$  est un produit scalaire hermitien sur  $\mathbb{C}$ .

Vu que  $p > 2$ , la fonction  $s: x \mapsto x^{p/2}$  est convexe (lemme 17.87). Calcul :

$$|f(x) - g(x)|^p + |f(x) + g(x)|^p = (|f(x) - g(x)|^2)^{p/2} + (|f(x) + g(x)|^2)^{p/2} \quad (27.667a)$$

$$= s(|\dots|^2) + s(|\dots|^2) \quad (27.667b)$$

$$\leq (|f(x) - g(x)|^2 + |f(x) + g(x)|^2)^{p/2} \quad (27.667c)$$

$$= (2|f(x)|^2 + 2|g(x)|^2)^{p/2} \quad (27.667d)$$

$$= 2^{p/2} (|f(x)|^2 + |g(x)|^2)^{p/2} \quad (27.667e)$$

$$\leq 2^{p/2} 2^{p/2-1} (|f(x)|^p + |g(x)|^p) \quad (27.667f)$$

$$= 2^{p-1} (|f(x)|^p + |g(x)|^p) \quad (27.667g)$$

Justifications :

- Pour (27.667c) : la convexité de  $s$ .
- Pour (27.667d) : la relation (27.666).
- Pour (27.667f) : la seconde inégalité du lemme 27.667f.

Nous isolons  $|f(x) - g(x)|^p$  :

$$|f(x) - g(x)|^p \leq 2^{p-1} (|f(x)|^p + |g(x)|^p) - |f(x) + g(x)|^p \quad (27.668a)$$

$$= 2^p \left( \frac{|f(x)|^p + |g(x)|^p}{2} - \left| \frac{|f(x)| + |g(x)|}{2} \right|^p \right) \quad (27.668b)$$

Cette inégalité étant valable pour tout  $x$ , nous pouvons intégrer sur  $\Omega$  et découper l'intégrale en petits morceaux :

$$\|f - g\|_p^p \leq 2^p \left( \frac{\|f\|_p^p + \|g\|_p^p}{2} - \left\| \frac{f + g}{2} \right\|_p^p \right). \quad (27.669)$$

Voilà une bonne chose de prouvée. Nous pouvons maintenant passer au vif du sujet.

Soit une suite  $w_j$  dans  $W$  telle que  $\|u_0 - w_j\| \rightarrow d(u_0, W)$ . Trois choses à savoir sur cette suite :

- (2a) Une telle suite existe parce que  $d(u_0, W)$  est défini comme un infimum.
- (2b) Rien ne garantit qu'elle converge.
- (2c) Même si elle convergeait, rien ne garantirait que la limite soit encore dans  $W$ .

Le troisième point est facile à régler : vu que  $W$  est fermé par hypothèse, une suite convergente contenue dans  $W$  a sa limite dans  $W$ . Nous allons régler la convergence de  $w_j$  en prouvant qu'elle est de Cauchy.

Remarquons que  $W$  est vectoriel, donc  $(w_j + w_k)/2$  est dans  $W$  pour tout  $j$  et  $k$  ; donc

$$\left\| \frac{w_j + w_k}{2} - u_0 \right\| \geq d(u_0, W). \quad (27.670)$$

En tenant compte de cela, nous écrivons l'inégalité (27.669) avec  $f = w_j - u_0$  et  $g = w_k - u_0$  :

$$\|f - g\|_p^p = \|w_j - w_k\|_p^p \leq 2^p \left( \frac{\|w_j - u_0\|^p + \|w_k - u_0\|^p}{2} - d(u_0, W)^p \right). \quad (27.671)$$

Soit  $\epsilon > 0$  et  $0 < \epsilon_1, \epsilon_2 < \epsilon$  tels que  $\epsilon_1 + \epsilon_2 < \epsilon$ . Il existe un  $N$  tel que si  $j, k > N$  alors  $\|w_j - u_0\|^p \leq d(u_0, W)^p + \epsilon_1$  et  $\|w_k - u_0\|^p \leq d(u_0, W)^p + \epsilon_2$ . Pour de telles valeurs de  $j$  et  $k$ , nous avons

$$\|w_j - w_k\|_p \leq 2 \left( \frac{\epsilon_1 + \epsilon_2}{2} \right)^{1/p} < 2\epsilon^{1/p}. \quad (27.672)$$

Donc la suite  $(w_j)$  est de Cauchy.

L'espace  $L^p$  étant complet par le théorème 27.55, nous en déduisons que  $(w_j)$  converge dans  $L^p$ . Mais comme  $W$  est fermé, nous avons  $w_j \xrightarrow{L^p} w \in W$ .

En termes de normes, nous avons

$$\|w - u_0\| = \lim_j \|w_j - u_0\| = d(W, u_0). \quad (27.673)$$

□

j

## 27.13 Jauge

DEFooRNZBooLSdowc

**Définition 27.156** ([259]).

Soit un ouvert connexe  $C$  dans un espace vectoriel topologique  $E$ . Nous supposons que  $0 \in C$ . La **jauge** de  $C$  est l'application

$$j_C: E \rightarrow \mathbb{R} \\ x \mapsto \inf\{t > 0 \text{ tel que } \frac{x}{t} \in C\}. \quad (27.674)$$

LEMooEUETooMAAnNO

**Lemme 27.157** ([1]).

Soit un ouvert connexe  $C$  dans un espace vectoriel topologique  $E$ . Nous supposons que  $0 \in C$ , et nous considérons la jauge  $j_C: E \rightarrow \mathbb{R}$ . Nous posons

$$A_x = \{t > 0 \text{ tel que } \frac{x}{t} \in C\}. \quad (27.675)$$

Nous avons :

ITEMooORQGooUkHWKJ

(1)  $]j_C(x), \infty[ \subset A_x$ .

ITEMooFZSGooLuyzcv

(2) Si  $x_0 \notin C$ , alors  $j_C(x_0) \geq 1$ .

*Démonstration.* Soit  $t \in A_x$ . Vu que  $C$  est connexe et contient 0, vu que  $x/t \in C$  nous avons

$$(1 - \lambda)0 + \lambda \frac{x}{t} \in C \quad (27.676)$$

pour tout  $\lambda \in [0, 1]$ . C'est à dire que  $(\lambda/t)x \in C$  pour tout  $\lambda \in [0, 1]$ . Donc  $(t/\lambda) \in A_x$ , et donc  $[t, \infty[ \subset A_x$ . Soit maintenant  $t > j_C(x)$ . Alors il existe  $t_0 \in [j_C(x), t[ \cap A_x$ . Dans ce cas,  $t \in [t_0, \infty[ \subset A_x$ . Cela prouve le point (1).

Nous passons au point (2). Soit  $x_0 \notin C$ . Si  $j_C(x_0) < 1$ , alors nous avons  $1 \in A_{x_0}$  par le point (1). Et si  $1 \in A_{x_0}$ , alors  $x_0 \in C$ , ce qui est contraire à l'hypothèse. Donc  $j_C(x_0) \leq 1$ . □

PROPooEBPPooDWUwLx

**Proposition 27.158** ([259]).

Soient un espace vectoriel topologique  $E$ , un ouvert connexe  $C \subset E$  tel que  $0 \in C$ . Nous la jauge  $j_C: E \rightarrow \mathbb{R}$  a les propriétés suivantes.

ITEMooAOUUooGhrpk

(1) Pour tout  $x \in E$  et  $\lambda > 0$  nous avons  $f_C(\lambda x) = \lambda j_C(x)$ .

ITEMooQWIQooMRramc

(2) Pour tout  $x, y \in E$  nous avons  $j_C(x + y) \leq j_C(x) + j_C(y)$ .

ITEMooKGNFfooKAEUPH

(3) Si  $C$  est symétrique<sup>77</sup>, alors  $j_C$  est une seminorme<sup>78</sup>

*Démonstration.* Nous posons

$$A_x = \{t > 0 \text{ tel que } \frac{x}{t} \in C\} \quad (27.677)$$

et  $A = A_x$ ,  $B = A_{\lambda x}$ .

(1)  $B = \lambda A$  Soit  $t \in B$ . Alors  $x/(t/\lambda) \in C$ . Et donc  $t/\lambda \in A$  et donc  $t \in \lambda A$ . De même si  $t \in A$ , alors  $x/t \in C$  et donc  $(\lambda x)/(\lambda t) \in C$  de telle sorte que  $\lambda t \in B$ .

(2) infimum, point (1) La proposition 1.569 dit que  $\inf(B) = \inf(\lambda A) = \lambda \inf(A)$ . Cela prouve le point (1).

77. C'est à dire que  $-x \in C$  si et seulement si  $x \in C$ .

78. Définition 7.380.

- (3) **Pour (3)** Nous savons par le lemme 27.157(1)  $]j_C(x), \infty[ \subset A_x$ . Si  $\epsilon > 0$ , nous avons donc  $\frac{x}{j_C(x) + \epsilon} \in A_x$  et donc

$$\frac{x}{j_C(x) + \epsilon} \in C, \quad \frac{y}{j_C(y) + \epsilon} \in C. \quad (27.678)$$

Comme  $C$  est convexe, nous pouvons prendre une combinaison convexe de ces deux éléments. L'astuce est d'écrire

$$\lambda x + (1 - \lambda)y \in C \quad (27.679)$$

avec

$$\lambda = \frac{j_C(x) + \epsilon}{j_C(x) + j_C(y) + 2\epsilon}, \quad (27.680)$$

et

$$1 - \lambda = \frac{j_C(y) + \epsilon}{j_C(x) + j_C(y) + 2\epsilon}. \quad (27.681)$$

Un simple calcul montre que

$$\frac{x + y}{j_C(x) + j_C(y) + 2\epsilon} \in C, \quad (27.682)$$

c'est-à-dire  $j_C(x) + j_C(y) + 2\epsilon \in A_{x+y}$  ou encore que

$$j_C(x + y) = \inf(A_{x+y}) \leq j_C(x) + j_C(y) + 2\epsilon. \quad (27.683)$$

Vu que cette inégalité est vraie pour tout  $\epsilon > 0$ , nous avons  $j_C(x + y) \leq j_C(x) + j_C(y)$ .

- (4) **Pour (3)** Nous devons prouver que  $j_C(\lambda x) = |\lambda|j_C(x)$  pour tout  $\lambda$ . Nous l'avons déjà prouvé pour  $\lambda > 0$ . Vu que  $0/t = 0 \in C$  pour tout  $t > 0$ , nous avons  $j_C(0) = 0$  et donc  $j_C(\lambda x) = |\lambda|j_C(x)$  pour  $\lambda = 0$ . Reste à voir cette égalité pour  $-\lambda$  lorsque  $\lambda > 0$ .

Vu que  $C$  est symétrique,  $A_x = A_{-x}$  et alors

$$j_C(-\lambda x) = \inf\{t > 0 \text{ tel que } \frac{-\lambda x}{t} \in C\} \quad (27.684a)$$

$$= \inf\{t > 0 \text{ tel que } \frac{\lambda x}{t} \in C\} \quad (27.684b)$$

$$= \lambda j_C(x). \quad (27.684c)$$

□

## 27.14 Théorèmes de Hahn-Banach

### 27.14.1 Applications $\mathbb{R}$ -linéaires et $\mathbb{C}$ -linéaires

**Lemme 27.159** ([759, 760]).

Soit un espace vectoriel  $X$  sur  $\mathbb{C}$ . Nous considérons<sup>79</sup> l'application

$$\begin{aligned} \psi: \mathcal{L}_{\mathbb{R}}(X, \mathbb{R}) &\rightarrow \mathcal{L}_{\mathbb{C}}(X, \mathbb{C}) \\ f &\mapsto [x \mapsto f(x) - if(ix)]. \end{aligned} \quad (27.685)$$

(1) Cette application est bien définie.

(2) Elle est une bijection.

(3) Elle vérifie  $\|\psi(f)\|_{\mathcal{L}_{\mathbb{C}}(X, \mathbb{C})} = \|f\|_{\mathcal{L}_{\mathbb{R}}(X, \mathbb{R})}$ .

(4) Si  $g \in \mathcal{L}_{\mathbb{C}}(X, \mathbb{C})$  se décompose en  $g(x) = u(x) + iv(x)$  où  $u$  et  $v$  sont à valeurs réelles, alors  $g = \psi(u)$ .

*Démonstration.* En plusieurs parties.

<sup>79</sup>. Voir la définition 4.31 pour les notations.

- (1) **Bien définie** Nous devons avant tout prouver que (27.685) a un sens : si  $f \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$ , alors nous devons prouver que  $\psi(f) \in \mathcal{L}_{\mathbb{C}}(X, \mathbb{C})$ , c'est-à-dire que  $\psi(f)$  est  $\mathbb{C}$ -linéaire. Soient donc  $x, y \in X$  et  $\alpha, \beta \in \mathbb{R}$  de telle sorte que  $\alpha + \beta i$  soit un élément générique de  $\mathbb{C}$ . Prouver que

$$\psi(f)(x + y) = \psi(f)(x) + \psi(f)(y) \quad (27.686)$$

est un simple calcul. Ensuite

$$\psi(f)((\alpha + i\beta)x) = f((\alpha + i\beta)x) - if(i(\alpha + i\beta)x) \quad (27.687a)$$

$$= \alpha f(x) + \beta f(ix) - i\alpha f(ix) - i(-\beta)f(x) \quad (27.687b)$$

$$= (\alpha + i\beta)(f(x) - if(ix)) \quad (27.687c)$$

$$= (\alpha + i\beta)\psi(f)(x). \quad (27.687d)$$

- (2)  **$\psi$  est injective** Soient  $f, g \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$  telles que  $\psi(f) = \psi(g)$ . Nous avons l'égalité

$$f(x) - if(ix) = g(x) - ig(ix), \quad (27.688)$$

qui est une égalité dans  $\mathbb{C}$ . En sachant que les nombres  $f(x)$ ,  $f(ix)$ ,  $g(x)$  et  $g(ix)$  sont des réels, nous séparons les parties réelles et imaginaires dans (27.688) et nous trouvons  $f(x) = g(x)$  et  $f(ix) = g(ix)$ . Chacune de ces deux égalités nous assurent que  $f = g$ .

- (3)  **$\psi$  est surjective** Soit  $g \in \mathcal{L}_{\mathbb{C}}(X, \mathbb{C})$ . Nous séparons ses parties réelles et imaginaires :

$$g(x) = g_1(x) + ig_2(x). \quad (27.689)$$

Avec  $g_1, g_2 \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$ . Nous allons prouver que  $g = \psi(g_1)$ .

Ensuite, utilisant la  $\mathbb{C}$ -linéarité de  $g$  pour calculer  $g(ix)$  de deux façon différentes. D'une part

$$g(ix) = g_1(ix) + ig_2(ix), \quad (27.690)$$

et d'autre part

$$g(ix) = i(g_1(x) + ig_2(x)). \quad (27.691)$$

En égalisant les parties réelles et imaginaires de (27.690) et de (27.691),

$$g_1(x) = g_2(ix) \quad (27.692a)$$

$$g_2(x) = -g_1(ix), \quad (27.692b)$$

et en particulier  $g_1(ix) = g_2(-x)$ .

Nous pouvons maintenant calculer

$$\psi(g_1)(x) = g_1(x) - ig_1(ix) \quad (27.693a)$$

$$= g_1(x) - ig_2(-x) \quad (27.693b)$$

$$= g_1(x) + ig_2(x) \quad (27.693c)$$

$$= g(x). \quad (27.693d)$$

- (4) **Norme** Nous prouvons maintenant le point (3) qui prétend que  $\|\psi(f)\|_{\mathcal{L}_{\mathbb{C}}(X, \mathbb{C})} = \|f\|_{\mathcal{L}_{\mathbb{R}}(X, \mathbb{R})}$ . Nous allons montrer les inégalités dans les deux sens.

$$(4a) \quad \|f\| \leq \|\psi(f)\|$$

D'abord nous prouvons que  $\|f\| \leq \|\psi(f)\|$ . Pour cela, nous considérons  $x \in X$ , et nous écrivons

$$|f(x)| \leq |f(x) - if(ix)| = |\psi(f)x| \quad (27.694)$$

dont la première inégalité est vraie parce que  $f(x)$  et  $f(ix)$  sont réels. Avec cette inégalité nous avons tout de suite

$$\sup_{\|x\|=1} |f(x)| \leq \sup_{\|x\|=1} |\psi(f)x|. \quad (27.695)$$

$$(4b) \quad \|f\| \geq \|\psi(f)\|$$

Nous passons maintenant à l'autre sens. Soit  $x \in X$ . Nous commençons par trouver  $\alpha \in S^1$  (c'est-à-dire  $|\alpha| = 1$ ) tel que  $f(\alpha ix) = 0$ . Si  $f(ix) = 0$ , alors  $\alpha = 1$  fait l'affaire. Sinon, supposons  $f(ix) \neq 0$ , et considérons l'application

$$\begin{aligned} s: ]0, 2\pi[ &\rightarrow \mathbb{R} \\ t &\mapsto f(\varphi(t)ix). \end{aligned} \quad (27.696)$$

où  $\varphi$  est la bijection continue de la proposition 18.62(1). D'une part, pour  $t = \pi$ , nous avons

$$s(\pi) = f(\varphi(\pi)ix) = f(-ix) = -f(ix), \quad (27.697)$$

et d'autre part, pour  $t = 0$  nous avons

$$s(0) = f(\varphi(0)ix) = f(ix) = f(ix). \quad (27.698)$$

Comme nous avons supposé que  $f(ix) \neq 0$  le théorème des valeurs intermédiaires 9.95 appliqué à  $s$  donne un  $t \in ]0, \pi[$  tel que  $s(t_0) = 0$ . En posant  $\alpha = \varphi(t_0)$  nous avons le  $\alpha$  qu'il nous fallait.

Armés de ce  $\alpha$ , nous montrons que  $\|\psi(f)\| \leq \|f\|$ . Nous avons

$$|\psi(f)x| = |\alpha| |\psi(f)x| \quad (27.699a)$$

$$= |\psi(f)(\alpha x)| \quad \psi(f) \text{ est } \mathbb{C}\text{-linéaire} \quad (27.699b)$$

$$= |f(\alpha x) - if(ix)| \quad (27.699c)$$

$$= |f(\alpha x)|. \quad (27.699d)$$

Nous avons prouvé que pour tout  $x \in X$  tel que  $\|x\| = 1$ , il existe  $y \in X$  tel que  $\|y\| = 1$  et  $|\psi(f)x| = |f(y)|$ . En l'occurrence,  $y = \alpha x$ . Notez que  $\alpha$  est une fonction de  $x$ . Nous avons donc

$$\sup_{\|y\|=1} |f(y)| \geq \sup_{\|x\|=1} |\psi(f)x|, \quad (27.700)$$

autrement dit  $\|f\| \geq \|\psi(f)\|$ .

- (5) **Décomposition** Nous prouvons le point (4). Supposons  $g(x) = u(x) + iv(x)$ . Nous avons  $g(ix) = ig(x)$  par  $\mathbb{C}$ -linéarité de  $g$ . Donc

$$u(ix) + iv(ix) = i(u(x) + iv(x)), \quad (27.701)$$

c'est-à-dire  $u(ix) + iv(ix) = iu(x) - v(x)$ . En égalisant les parties imaginaires,  $v(x) = -u(ix)$ . En remettant dans l'expression de  $g$ ,

$$g(x) = u(x) + iv(x) = u(x) + i(-u(ix)) = u(x) - iu(ix) = \psi(u)x, \quad (27.702)$$

ce qu'il fallait démontrer. □

LEMooUFM FooEXecXE

**Lemme 27.160** ([759]).

Soient un espace vectoriel  $X$  sur  $\mathbb{C}$  ainsi qu'une seminorme<sup>80</sup>  $p$ . Soit  $f \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$ . Nous avons

$$|\psi(f)(x)| \leq p(x) \quad \forall x \in X \quad (27.703)$$

si et seulement si

$$|f(x)| \leq p(x) \quad \forall x \in X \quad (27.704)$$

*Démonstration.* En deux parties.

---

<sup>80</sup>. Définition 7.380.

- (1)  $\Rightarrow$  Nous supposons que  $f \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$  vérifie  $|\psi(f)(x)| \leq p(x)$  pour tout  $x$ . L'observation à faire est  $f(x) = \operatorname{Re}(\psi(f)(x))$ . Nous avons alors<sup>81</sup>

$$|f(x)| = |\operatorname{Re}(\psi(f)(x))| \leq |\psi(f)(x)| \leq p(x). \quad (27.705)$$

- (2)  $\Leftarrow$  Nous supposons que  $|f(x)| \leq p(x)$  pour tout  $x$ . Soit  $x \in X$ . Si  $\psi(f)(x) = 0$ , alors nous avons bien  $|\psi(f)(x)| \leq p(x)$  parce qu'une seminorme est toujours positive. Nous supposons donc que  $\psi(f)(x) \neq 0$ .

La décomposition polaire (proposition 18.63) du nombre complexe  $\psi(f)x$  donne un  $\theta \in \mathbb{R}$  tel que  $\psi(f)x = e^{i\theta}|\psi(f)x|$ . Donc

$$|\psi(f)(x)| = e^{-i\theta}\psi(f)(x) = \psi(f)(e^{-i\theta}x). \quad (27.706)$$

Ces égalités montrent entre autres que  $\psi(f)(e^{-i\theta}x) \in \mathbb{R}^+$  ; et en particulier, il est égal à sa partie réelle. Vu la définition (27.685), ça nous dit que

$$\psi(f)(e^{-i\theta}x) = f(e^{-i\theta}x). \quad (27.707)$$

Nous pouvons alors continuer les égalités (27.706) en mettant des normes partout :

$$|\psi(f)(x)| = |\psi(f)(e^{-i\theta}x)| = |f(e^{-i\theta}x)| \leq p(e^{-i\theta}x) = |e^{-i\theta}|p(x) = p(x). \quad (27.708)$$

□

### 27.14.2 Hahn-Banach, théorème d'extension dominée

**Théorème 27.161** (Hahn-Banach, extension dominée cas réel[761, 259]).

Soit  $E$ , un espace vectoriel réel et une application  $p: E \rightarrow \mathbb{R}$  satisfaisant<sup>82</sup>

- (1)  $p(\lambda x) = \lambda p(x)$  pour tout  $x \in E$  et pour tout  $\lambda > 0$ ,
- (2)  $p(x + y) \leq p(x) + p(y)$  pour tout  $x, y \in E$ .

Soit de plus  $M \subset E$  un sous-espace vectoriel muni d'une application  $g: M \rightarrow \mathbb{R}$  vérifiant  $g(m) \leq p(m)$  pour tout  $m \in M$ . Alors il existe  $f \in \mathcal{L}(E, \mathbb{R})$  telle que

- (1) L'application  $f$  étend  $g$  :  $f(m) = g(m)$  pour tout  $m \in M$ ,
- (2) L'application  $f$  reste dominée par  $p$  :  $f(x) \leq p(x)$  pour tout  $x \in E$ .

*Démonstration.* Si  $h$  est une application linéaire définie sur un sous-espace de  $E$ , nous notons  $D_h$  ledit sous-espace.

#### (1) Un ensemble inductif

Nous considérons  $P$ , l'ensemble des fonctions linéaires suivant

$$P = \left\{ h: D_h \rightarrow \mathbb{R} \text{ tel que } \begin{cases} M \subset D_h \\ h(m) = g(m) \quad \forall m \in M \\ h(x) \leq p(x) \quad \forall x \in D_h \end{cases} \right\} \quad (27.709)$$

Cet ensemble est non vide parce que  $g$  est dedans. Nous le munissons de la relation d'ordre  $h_1 \leq h_2$  si et seulement si  $D_{h_1} \subset D_{h_2}$  et  $h_2$  prolonge  $h_1$ . Nous montrons à présent que  $P$  est un ensemble inductif. Soit un sous-ensemble totalement ordonné  $Q \subset P$  ; nous définissons une fonction  $h$  de la façon suivante. D'abord  $D_h = \sup_{l \in Q} D_l$  et ensuite

$$\begin{aligned} h: D_h &\rightarrow \mathbb{R} \\ x &\mapsto l(x) \quad \text{si } x \in D_l \end{aligned} \quad (27.710)$$

81. Dans [759], l'auteur sépare le calcul en deux parties : une majoration pour  $f(x)$  et une pour  $-f(x)$ . Je ne vois pas où est le mal à la faire d'un seul coup.

82. Ce n'est pas tout à fait une seminorme (définition 7.380) à cause de la valeur absolue. Le théorème 27.163 parlera pour les vraies seminormes, et pour les espaces sur  $\mathbb{C}$ .



Cela est bien défini parce que si  $x \in D_l \cap D_{l'}$  alors, vu que  $Q$  est totalement ordonné (i.e.  $l \leq l'$  ou  $l' \leq l$ ), on a obligatoirement  $D_l \subset D_{l'}$  et  $l'$  qui prolonge  $l$  (ou le contraire). Donc  $h$  est un majorant de  $Q$  dans  $P$  parce que  $h \geq l$  pour tout  $l \in Q$ . Cela montre que  $P$  est inductif (définition 1.36). Le lemme de Zorn 1.37 nous dit alors que  $P$  possède un élément maximal  $f$  qui va être la réponse à notre théorème.

(2) **Le support de  $f$**

La fonction  $f$  est dans  $P$ ; donc  $f(x) \leq p(x)$  pour tout  $x \in D_h$  et  $f(m) = g(m)$  pour tout  $m \in M$ . Pour terminer nous devons montrer que  $D_f = E$ . Supposons donc que  $D_f \neq E$  et prenons  $x_0 \notin D_f$ . Nous allons contredire la maximalité de  $f$  en considérant la fonction  $h$  donnée par  $D_h = D_f + \mathbb{R}x_0$  et

$$h(x + tx_0) = f(x) + t\alpha \quad (27.711)$$

où  $\alpha$  est une constante que nous allons fixer plus tard.

Nous commençons par prouver que  $h$  est dans  $P$ . Nous devons prouver que

$$h(x + tx_0) = f(x) + t\alpha \leq p(x + tx_0) \quad (27.712)$$

Pour cela nous allons commencer par fixer  $\alpha$  pour avoir les relations suivantes :

$$\begin{cases} f(x) + \alpha \leq p(x + x_0) \\ f(x) - \alpha \leq p(x - x_0) \end{cases} \quad (27.713a)$$

$$\quad (27.713b)$$

pour tout  $x \in D_f$ . Ces relations sont équivalentes à demander  $\alpha$  tel que

$$\begin{cases} \alpha \leq p(x + x_0) - f(x) \\ \alpha \geq f(x) - p(x - x_0) \end{cases} \quad (27.714a)$$

$$\quad (27.714b)$$

Nous nous demandons donc si il existe un  $\alpha$  qui satisfasse

$$\sup_{y \in D_f} (f(y) - p(y - x_0)) \leq \alpha \leq \inf_{z \in D_f} (p(z + x_0) - f(z)). \quad (27.715)$$

Ou encore nous devons prouver que pour tout  $y, z \in D_f$ ,

$$p(z + x_0) - f(x) \geq f(y) - p(y - x_0) \geq 0. \quad (27.716)$$

Par les propriétés de  $p$  et de  $f$ ,

$$p(z + x_0) + p(y - x_0) - f(z) - f(y) \geq p(z + y) - f(z + y) \geq 0. \quad (27.717)$$

La dernière inégalité est le fait que  $f \in P$ . Un choix de  $\alpha$  donnant les inéquations (27.713) est donc possible.

À partir des inéquations (27.713) nous obtenons la relation (27.712) de la façon suivante. Si  $t > 0$  nous multiplions l'équation (27.713a) par  $t$  :

$$tf(x) + t\alpha \leq tp(x + x_0). \quad (27.718)$$

Et nous écrivons cette relation avec  $x/t$  au lieu de  $x$  en tenant compte de la linéarité de  $f$  :

$$f(x) + t\alpha \leq tp\left(\frac{x}{t} + x_0\right) = p(x + tx_0). \quad (27.719)$$

Avec  $t < 0$ , c'est similaire, en faisant attention au sens des inégalités.

Nous avons donc construit  $h: D_h \rightarrow \mathbb{R}$  avec  $h \in P$ ,  $D_f \subset D_h$  et  $h(x) = f(x)$  pour tout  $x \in D_f$ . Cela pour dire que  $h > f$ , ce qui contredit la maximalité de  $f$ . Le domaine de  $f$  est donc  $E$  tout entier.

La fonction  $f$  est donc une fonction qui remplit les conditions.

□

**Lemme 27.162** ([296]).

Soient un espace vectoriel complexe  $X$ , et une application linéaire  $f: X \rightarrow \mathbb{C}$ . Il existe une application  $u \in \mathcal{L}_{\mathbb{R}}(X, \mathbb{R})$  telle que

ITEMooBWQUooRbPZSn

(1) Pour tout  $x \in X$ , nous avons

$$f(x) = u(x) - iu(ix) \quad (27.720)$$

ITEMooFIZBooDnYKSO

(2) Nous avons égalité des normes<sup>83</sup>

$$\|f\|_{\mathcal{L}_{\mathbb{C}}(X, \mathbb{C})} = \|u\|_{\mathcal{L}_{\mathbb{R}}(X, \mathbb{R})}. \quad (27.721)$$

*Démonstration.* En deux parties.(1) **Pour (1)**

Nous décomposons, pour chaque  $x \in X$ , l'élément  $f(x)$  en parties réelles et imaginaires :

$$f(x) = u(x) + iv(x). \quad (27.722)$$

EQooZJZJooBQxiffg

Il ne reste plus qu'à trouver le lien entre  $u$  et  $v$ . Pour cela, nous écrivons  $f(ix) = if(x)$  en substituant  $f(x)$  et  $f(ix)$  par leurs valeurs en termes de  $u$  et  $v$  :

$$u(ix) + iv(ix) = iu(x) - v(x). \quad (27.723)$$

En égalisant les parties réelles, nous avons  $u(ix) = -v(x)$  et donc bien

$$f(x) = u(x) + iv(x) = u(x) - iu(ix). \quad (27.724)$$

EQooZUTLooVzwRDM

Nous prouvons maintenant que  $u$  est  $\mathbb{R}$ -linéaire. Il suffit de considérer  $\lambda \in \mathbb{R}$  et d'écrire  $f(\lambda x) = \lambda f(x)$  en substituant (27.722) pour  $f(\lambda x)$  et  $f(x)$  :

$$u(\lambda x) + iv(\lambda x) = \lambda(u(x) + iv(x)). \quad (27.725)$$

En égalisant encore les parties réelles, nous avons  $u(\lambda x) = \lambda u(x)$ . Nous faisons de même pour prouver que  $u(x + y) = u(x) + u(y)$ .

(2) **Pour (2)**

□

THOooVQLJooWuBMoZ

**Théorème 27.163** (Hahn-Banach complexe[296]).

Soient un espace vectoriel  $X$  sur  $\mathbb{K}$  (qui est  $\mathbb{R}$  ou  $\mathbb{C}$ ), et une seminorme  $p: X \rightarrow \mathbb{R}$ . Soient un sous-espace vectoriel  $M$  de  $X$  ainsi qu'une application linéaire  $\ell \in \mathcal{L}(M, \mathbb{K})$  telle que

$$|\ell(m)| \leq p(m) \quad \forall m \in M. \quad (27.726)$$

Alors il existe une extension linéaire  $f \in \mathcal{L}(X, \mathbb{K})$  telle que

(1)  $f(m) = \ell(m)$  pour tout  $m \in M$ ,(2)  $|f(x)| \leq p(x)$  pour tout  $x \in X$ .

*Démonstration.* Nous considérons  $X$  comme espace vectoriel réel (de dimension double de celle de  $X$  comme espace vectoriel complexe). Là-dedans, l'application  $p$  vérifie  $p(\lambda x) = \lambda p(x)$  pour tout  $\lambda > 0$ . D'autre part nous décomposons l'application  $\ell: M \rightarrow \mathbb{C}$  en parties réelles et imaginaires :

$$\ell(x) = u(x) + iv(x), \quad (27.727)$$

où  $u: M \rightarrow \mathbb{R}$  est une application vérifiant, pour tout  $m \in M$ ,

$$u(m) \leq |u(m)| \leq |\ell(m)| \leq p(m). \quad (27.728)$$

---

83. Pas encore démontrée ; mais si vous savez le faire, écrivez-moi.

Donc le théorème en version réelle 27.161 s'applique et il existe une extension  $R: X \rightarrow \mathbb{R}$  de  $u$  qui vérifie  $R \leq p$  sur  $X$ . Nous posons alors

$$f(x) = R(x) - iR(ix). \quad (27.729)$$

En vertu du lemme 27.159, cela est une application  $\mathbb{C}$ -linéaire sur  $X$ , et nous avons  $f = \ell$  sur  $M$  parce que leurs parties réelles sont égales.

De plus nous avons  $R \leq p$  et donc  $|f| \leq p$  par le lemme 7.384.  $\square$

### 27.14.3 Hyperplan séparateur

DEFooQUDEooMrgcCV

**Définition 27.164** (Hyperplan qui sépare).

Soit  $E$  un espace vectoriel topologique réel ainsi que  $A, B$  des sous-ensembles de  $E$ .

- (1) Nous disons que l'hyperplan d'équation  $f = \alpha$  **sépare au sens large** les parties  $A$  et  $B$  si  $f(x) \leq \alpha$  pour tout  $x \in A$  et  $f(x) \geq \alpha$  pour tout  $x \in B$ .
- (2) La séparation est **au sens strict** si il existe  $\epsilon > 0$  tel que

$$f(x) \leq \alpha - \epsilon \quad \text{pour tout } x \in A \quad (27.730a)$$

$$f(x) \geq \alpha + \epsilon \quad \text{pour tout } x \in B. \quad (27.730b)$$

Si  $X$  est un espace vectoriel topologique complexe, nous disons que  $A$  et  $B$  sont **strictement séparés** si il existe  $\phi \in X'$  et  $\alpha > 0$  tels que

$$\operatorname{Re}(\phi)(a) > \alpha > \operatorname{Re}(\phi)(b) \quad (27.731)$$

pour tout  $a \in A$  et  $b \in B$ .

PROPooSKBXooQDmcl

**Proposition 27.165** ([259]).

Soit un espace vectoriel topologique réel  $E$ . Soit un ouvert connexe non vide  $C \subset E$ . Soit  $x_0 \notin C$ . Alors il existe  $f \in E'$ <sup>84</sup> tel que  $f(x) < f(x_0)$  pour tout  $x \in C$ .

*Démonstration.* En deux parties.

- (1) **Si  $0 \in C$**  Nous commençons par supposer que  $0 \in C$ , de telle sorte à pouvoir considérer la jauge<sup>85</sup>  $j_C: E \rightarrow \mathbb{R}$ . Nous posons  $M = \mathbb{R}x_0$  et

$$\begin{aligned} g: M &\rightarrow \mathbb{R} \\ tx_0 &\mapsto t. \end{aligned} \quad (27.732)$$

Vu que  $x_0 \notin C$ , le lemme 27.157(2) dit que  $j_C(x_0) \geq 1 = g(x_0)$ . Nous multiplions cette inégalité par  $t \geq 0$  pour obtenir  $tj_C(x_0) \geq tg(x_0)$ . Vu que  $tj_C(x_0) = j_C(tx_0)$  (proposition 27.158(1)), et que  $tg(x_0) = g(tx_0)$ , nous avons

$$j_C(tx_0) \geq g(tx_0) \quad (27.733)$$

EQooYSVTooFKJabg

pour tout  $t \geq 0$ . Si  $t < 0$ , alors nous avons  $j_C(tx_0) \geq 0$  et  $g(tx_0) = t < 0$ . Donc l'inégalité (27.733) est valide pour tout  $t \in \mathbb{R}$ . Nous avons donc  $j_C \geq g$  sur  $M$ . Le théorème 27.161 dit qu'il existe  $f \in \mathcal{L}(E, \mathbb{R})$  tel que

(1a)  $f(x) = g(x)$  pour tout  $x \in M$ .

(1b)  $f(x) \leq j_C(x)$  pour tout  $x \in E$ .

Pour  $x \in C$  en particulier nous avons  $f(x) \leq j_C(x) < 1 = f(x_0)$ . Nous devons encore prouver que  $f$  est continue. Pour cela nous considérons l'ouvert  $U = C \cap (-C)$ , qui est un voisinage ouvert de 0. Soit  $x \in U$ . Nous avons  $f(x) \leq j_C(x) < 1$  parce que  $x \in C$ , mais aussi  $f(-x) \leq j_C(-x) < 1$  parce que  $-x \in C$ . Vu que  $f$  est linéaire, l'inégalité  $f(-x) < 1$  donne  $f(x) > -1$ . Au final nous avons prouvé que

$$f(x) \in ]-1, 1[ \quad (27.734)$$

pour tout  $x \in U$ . La proposition 7.279 dit qu'alors  $f$  est continue.

84.  $E'$  est l'ensemble des applications linéaires continues sur  $E$  à valeurs dans le corps, c'est-à-dire  $\mathbb{R}$  ici.

85. Définition 27.156.

- (2) Si  $0 \notin C$  Enfin nous prouvons le cas où  $0$  n'est pas dans  $C$ . Soit  $c \in C$ . Nous posons  $C' = C - c$  et  $x'_0 = x_0 - c$  de telle sorte que  $0 \in C'$  et  $x'_0 \in C'$ . Par la première partie, il existe  $f \in E'$  telle que  $f(x) < f(x'_0)$  pour tout  $x \in C'$ . Nous posons ensuite

$$f_1(x) = f(x - c). \quad (27.735)$$

Nous avons alors, pour tout  $x \in C$  que

$$f_1(x) = f(x - c) < f(x'_0) = f(x_0 - c) = f_1(x_0). \quad (27.736)$$

parce que  $x - c \in C'$ .

L'application  $f_1$  est encore continue et linéaire. Elle répond donc à la question.

□

ThoSAJjdZc

**Théorème 27.166** (Hahn-Banach, première forme géométrique[259]).

Soit  $E$  un espace vectoriel topologique et  $A, B$  deux convexes non vides disjoints de  $E$ . Si  $A$  est ouvert, il existe un hyperplan fermé qui sépare  $A$  et  $B$  au sens large<sup>86</sup>.

*Démonstration.* Vu que  $A$  et  $B$  sont convexes, la proposition 7.210 dit que  $A - B$  est convexe. De plus  $A - B$  est ouvert et ne contient pas zéro. Nous pouvons donc lui appliquer la proposition 27.165 : il existe  $f \in E'$  tel que  $f(x) < f(0) = 0$  pour tout  $x \in A - B$ . Autrement dit, pour tout  $a, b \in A \times B$  nous avons  $f(a) < f(b)$ . Cela implique

$$\sup\{f(a) \text{ tel que } a \in A\} \leq \inf\{f(b) \text{ tel que } b \in B\}. \quad (27.737)$$

En posant  $\alpha \in \mathbb{R}$  tel que

$$\sup\{f|_A\} \leq \alpha \leq \inf\{f|_B\}, \quad (27.738)$$

l'hyperplan  $\{f = \alpha\}$  sépare bien  $A$  et  $B$  au sens large.

□

THOooEOQVvAdPdRgW

**Théorème 27.167** (Hahn-Banach, seconde forme géométrique[762]).

Soient un espace vectoriel topologique réel ou complexe localement convexe<sup>87</sup>  $X$  ainsi que des parties  $A, B$  qui sont

- (1) des convexes
- (2) non vides
- (3) disjoints
- (4)  $A$  compact et  $B$  soit fermé.

Alors il existe un hyperplan fermé qui sépare strictement<sup>88</sup>  $A$  et  $B$ .

Dans le cas complexe, cela signifie qu'il existe  $\phi \in X'$  et  $\alpha > 0$  tels que

$$\operatorname{Re}(\phi)(a) > \alpha > \operatorname{Re}(\phi)(b) \quad (27.739)$$

EQooTYESooXLmkE

pour tout  $a \in A$  et  $b \in B$ .

*Démonstration.* Nous séparons le cas réel du cas complexe.

(1) **Cas réel**

Vu que  $B$  est fermé,  $A$  est dans l'ouvert  $X \setminus B$ . Donc si  $a \in A$ , il existe un voisinage ouvert convexe de  $a$  inclus dans  $X \setminus B$ . Soit  $U_a$  un voisinage ouvert et convexe de  $0$  tel que  $(a + U_a) \cap B = \emptyset$ .

Vu que la fonction  $(x, y) \mapsto x + y$  est continue, nous pouvons trouver un ouvert convexe  $V_a$  tel que  $V_a + V_a \subset U_a$ . L'ensemble  $a + V_a$  est alors un voisinage ouvert de  $a$  et bien

86. Définition 27.164.

87. Définition 7.248.

88. Hyperplan strictement séparateur, définition 27.164.

entendu  $\bigcup_a (a + V_a)$  recouvre  $A$  qui est compact. Nous en extrayons un sous-recouvrement fini, c'est-à-dire que nous considérons  $a_1, \dots, a_n \in A$  tels que

$$A \subset \bigcup_{i=1}^n (a_i + V_{a_i}). \quad (27.740)$$

Nous posons alors

$$V = \bigcap_{i=1}^n V_{a_i}. \quad (27.741)$$

Cet ensemble est non vide parce qu'il contient un voisinage de zéro parce que c'est une intersection finie de voisinages de zéro. Soit  $x \in A + V$ . Il existe  $i$  tel que

$$x \in a_i + U_{a_i} + V \subset a_i + V_{a_i} + V_{a_i} \subset a_i + U_{a_i} \subset X \setminus B. \quad (27.742)$$

Donc  $(A + V) \cap B = \emptyset$ . L'ensemble  $A + V$  est alors un ouvert convexe disjoint de  $B$ . Par la première forme géométrique du théorème de Hahn-Banach 27.166 nous avons un hyperplan qui sépare  $A + V$  de  $B$  au sens large : il existe  $f \in X' \setminus \{0\}$  tel que  $f(a) + f(v) \leq f(b)$  pour tout  $a \in A$ ,  $v \in V$  et  $b \in B$ .

Il suffit donc de trouver un  $v \in V$  tel que  $f(v) \neq 0$  pour avoir la séparation au sens strict. Cela est facile :  $V$  étant un voisinage de zéro et  $f$  étant linéaire, si elle était nulle sur  $V$ , elle serait nulle sur  $X$ .

## (2) Cas complexe

L'espace  $X$  est complexe. Nous considérons  $E$  l'espace réel correspondant (proposition 7.224). Nous lui appliquons le cas réel : il existe  $\alpha \in \mathbb{R}$  et une application linéaire  $f: X \rightarrow \mathbb{R}$  telle que

$$f(a) < \alpha < f(b) \quad (27.743)$$

pour tout  $a \in A$  et  $b \in B$ . Notez que  $f \in E'$  et donc  $\|f\| < \infty$  (bornée si et seulement si continue, proposition 10.81). Nous considérons maintenant l'application linéaire

$$\phi(x) = f(x) - if(ix) \quad (27.744)$$

décrite dans la proposition 7.225. C'est une application continue parce que  $\|\phi\| = \|f\| < \infty$ . De plus  $\operatorname{Re}(\phi) = f$  et  $\phi$  vérifie donc les inégalités 27.739.

□

COR00HTZVooFhgrSN

### Corolaire 27.168 ([259]).

Soit un espace vectoriel réel localement convexe. Soit un sous-espace vectoriel  $F$  tel que  $\bar{F} \neq E$ . Alors il existe une application  $f \in E' \setminus \{0\}$  telle que  $f = 0$  sur  $F$ .

*Démonstration.* Soit  $a \in E$  avec  $a \notin \bar{F}$ . Vu que le singleton  $\{a\}$  est compact et que  $\bar{F}$  est fermé, le théorème de Hahn-Banach 27.167 dit qu'il existe  $f \in E' \setminus \{0\}$  et  $\alpha \neq 0$  tel que  $f < \alpha$  sur  $\bar{F}$  et  $f(a) > \alpha$ .

Étant donné que  $F$  est vectoriel et que  $f$  est linéaire, avoir  $f(x) < \alpha$  pour tout  $x$  dans  $F$  implique d'avoir  $f(x) = 0$  pour tout  $x \in F$ . En effet si  $f(x) \neq 0$ , il existe un  $\lambda$  tel que  $f(\lambda x) > \alpha$  ( $\mathbb{R}$  est archimédien, théorème 1.543). □

## 27.14.4 Prolongement de fonctionnelles (dimension finie)

Nous allons prouver quelques lemmes qui permettent de prolonger des fonctionnelles d'un sous-espace vers un sous-espace contenant un nombre fini de dimensions en plus.

LEM00HWSJooGVmIPV

### Lemme 27.169 ([763]).

Soient un espace vectoriel réel normé  $X$  ainsi qu'un sous-espace  $M$ . Soient  $\ell \in \mathcal{L}(M, \mathbb{R})$  de norme finie, et  $x_1 \in X \setminus M$ . On pose  $M_1 = \operatorname{Span}\{M, x_1\}$ . Alors il existe  $\ell_1 \in \mathcal{L}(M_1, \mathbb{R})$  tel que

(1)  $\ell_1(x) = \ell(x)$  pour tout  $x \in M$ ,

(2)  $\|\ell_1\|_{\mathcal{L}(M_1, \mathbb{R})} = \|\ell\|_{\mathcal{L}(M, \mathbb{R})}$

*Démonstration.* Si  $\ell = 0$ , alors c'est facile : on prend  $\ell_1 = 0$ . Nous commençons par supposer que  $\|\ell\| = 1$  ; nous ferons le cas général ensuite.

(1) **Deux fonctions** Nous considérons les fonctions

$$\begin{aligned} f_+ : M &\rightarrow \mathbb{R} \\ z &\mapsto \|x_1 + z\| - \ell(z) \end{aligned} \quad (27.745)$$

et

$$\begin{aligned} f_- : M &\rightarrow \mathbb{R} \\ z &\mapsto -\|x_1 + z\| - \ell(z). \end{aligned} \quad (27.746)$$

Pour tout  $z_1, z_2 \in M$  nous avons

$$f_+(z_1) - f_-(z_2) = \|x_1 + z_1\| - \ell(z_1) + \|x_1 + z_2\| + \ell(z_2) \quad (27.747a)$$

$$= \|x_1 + z_1\| + \|x_1 + z_2\| - \ell(z_1 - z_2) \quad (27.747b)$$

$$\geq \|(x_1 + z_1) - (x_1 + z_2)\| - \ell(z_1 - z_2) \quad (27.747c)$$

$$= \|z_1 - z_2\| - \ell(z_1 - z_2) \quad (27.747d)$$

$$\geq 0 \quad (27.747e)$$

parce que  $\|\ell\| = 1$ . Nous en déduisons que  $f_+(z_1) \geq f_-(z_2)$ .

(2) **Un entre les deux**

En posant

$$c_- = \sup_{z \in M} (-\|x_1 + z\| - \ell(z)) \quad (27.748a)$$

$$c_+ = \inf_{z \in M} (\|x_1 + z\| - \ell(z)), \quad (27.748b)$$

nous avons  $c_- \leq c_+$ . Nous choisissons  $c_1 \in [c_-, c_+]$ .

(3) **La définition** Si  $x \in M_1 = \text{Span } M, x_1$ , alors il existe  $y \in M$  et  $\lambda \in \mathbb{R}$  tels que

$$x = \lambda x_1 + y. \quad (27.749)$$

Nous posons alors

$$\ell_1(x) = \lambda c_1 + \ell(y). \quad (27.750)$$

Voilà qui définit notre  $\ell_1$ . Nous devons prouver qu'elle satisfait  $\|\ell_1\| = 1$  et  $\ell_1(x) = \ell(x)$  pour tout  $x \in M$ . La seconde condition est facile : si  $x \in M$ , alors  $\lambda = 0$  dans (27.750) et nous avons bien  $\ell_1(x) = \ell(x)$ .

Pour la condition sur la norme, nous allons devoir un peu travailler.

(4)  $|\ell_1(x)| \leq \|x\|$  pour tout  $x \in M_1$  Soit  $x \in M_1$ . Nous avons  $x = \lambda x_1 + y$  avec  $\lambda \in \mathbb{R}$  et  $y \in M$ .  
i

(4a) **Si  $\lambda = 0$**  Alors  $x \in M$  et  $|\ell_1(x)| = |\ell(x)| \leq \|x\|$  parce que  $\|\ell\| = 1$ .

(4b) **Si  $\lambda \neq 0$**  Soient  $z_1, z_2 \in M$ . Par définition de  $c_1, c_+$  et  $c_-$  nous avons les inégalités

$$-\|x_1 + z_1\| - \ell(z_1) \leq c_- \leq c_1 \leq c_+ \leq \|x_1 + z_2\| - \ell(z_2). \quad (27.751)$$

Nous écrivons ces inégalités pour  $z_1 = z_2 = y/\lambda$  :

$$-\|x_1 + \frac{y}{\lambda}\| - \ell(y/\lambda) \leq c_1 \leq \|x_1 + \frac{y}{\lambda}\| - \ell(y/\lambda). \quad (27.752)$$

(4c) Si  $\lambda > 0$  Nous multiplions (27.752) par  $\lambda$  et nous profitons de la linéarité de  $\ell$  :

$$-\|\lambda x_1 + y\| - \ell(y) \leq \lambda c_1 \leq \|\lambda x_1 + y\| + \ell(y), \quad (27.753)$$

donc

$$-\|\lambda x_1 + y\| \leq \underbrace{\lambda c_1 + \ell(y)}_{=\ell_1(\lambda x_1 + y)} \leq \|\lambda x_1 + y\| \quad (27.754)$$

Nous en déduisons que

$$|\ell(\lambda x_1 + y)| \leq \|\lambda x_1 + y\|, \quad (27.755)$$

ce qu'il fallait.

(4d) Si  $\lambda < 0$  Le calcul est le même, mais il faut faire attention à bien reverser les inégalités au bon moment, et en manipulant bien les valeurs absolues. Nous avons par exemple

$$\lambda \|x_1 + \frac{y}{\lambda}\| = -\| |\lambda| x_1 + \frac{|\lambda| y}{\lambda} \| = -\| -\lambda x_1 - y \| = -\|\lambda x_1 + y\|. \quad (27.756)$$

En multipliant encore (27.752) par  $\lambda$ , nous trouvons

$$-\lambda \|x_1 + \frac{y}{\lambda}\| - \ell(y) \geq \lambda c_1 \geq \lambda \|x_1 + \frac{y}{\lambda}\| - \ell(y). \quad (27.757)$$

qui devient

$$\|\lambda x_1 + y\| \geq \lambda c_1 + \ell(y) \geq -\|\lambda x_1 + y\|, \quad (27.758)$$

qui revient au même que (27.754).

(5) **Première conclusion** Nous avons prouvé que  $|\ell_1(x)| \leq \|x\|$  pour tout  $x$  dans  $M_1$ . Cela signifie que  $\|\ell_1\|_{M_1'} \leq 1$ . Pour prouver que  $\|\ell_1\| = 1$  nous prouvons l'inégalité inverse :

$$\|\ell_1\|_{M_1'} = \sup_{x \in M_1} \frac{|\ell_1(x)|}{\|x\|} \geq \sup_{x \in M} \frac{|\ell_1(x)|}{\|x\|} = \sup_{x \in M} \frac{|\ell(x)|}{\|x\|} = \|\ell\|_M = 1. \quad (27.759)$$

Nous en déduisons que  $\|\ell_1\| = 1$  et cela termine la preuve dans le cas où  $\|\ell\| = 1$ .

Maintenant, si  $\|\ell\| = a \neq 1$ , nous considérons la forme linéaire  $f = \ell/a$  qui satisfait  $\|f\| = 1$ . Par la partie déjà prouvée, nous définissons une extension  $f_1 : M_1 \rightarrow \mathbb{R}$  telle que  $\|f_1\| = 1$ .

Il suffit alors de poser  $\ell_1 = a f_1$ , et nous avons le résultat.  $\square$

Ce lemme est également valable pour les complexes.

**Lemme 27.170** ([1]).

Soit un espace vectoriel normé  $X$  sur  $\mathbb{C}$ . Soit  $\ell \in \mathcal{L}_{\mathbb{C}}(X, \mathbb{C})$  de norme finie. Soient un sous-espace  $M$  de  $X$  ainsi qu'un élément  $a \in X \setminus M$ . Il existe  $\ell_1 \in \mathcal{L}_{\mathbb{C}}(\text{Span}_{\mathbb{C}}(M, a), \mathbb{C})$  tel que

(1)  $\ell_1(x) = \ell(x)$  pour tout  $x \in M$ ,

(2)  $\|\ell_1\| = \|\ell\|$ .

*Démonstration.* Nous considérons l'application  $\psi$  du lemme 27.159. En posant  $f = \psi^{-1}(\ell)$ , nous avons  $f \in \mathcal{L}_{\mathbb{R}}(M, \mathbb{R})$  vérifiant  $\|\ell\| = \|f\|$ .

L'espace  $X$  peut être vu comme vectoriel sur  $\mathbb{R}$ . Le lemme 27.169 permet de prolonger  $f$  à

$$f_1 \in \mathcal{L}_{\mathbb{R}}(\text{Span}_{\mathbb{R}}(M, a), \mathbb{R}) \quad (27.760)$$

vérifiant  $\|\ell\| = \|f\| = \|f_1\|$ .

Vu que  $M$  est déjà un espace vectoriel sur  $\mathbb{C}$ , l'espace  $\text{Span}_{\mathbb{C}}(M, a)$  qui nous intéresse est donné par

$$\text{Span}_{\mathbb{C}}(M, a) = \text{Span}_{\mathbb{R}}(M, a, ia). \quad (27.761)$$

Nous pouvons donc utiliser une deuxième fois le lemme 27.169 avec le vecteur  $ia$ , et définir une extension

$$f_2 \in \mathcal{L}_{\mathbb{R}}(\text{Span}_{\mathbb{R}}(M, a, ia), \mathbb{R}) = \mathcal{L}_{\mathbb{R}}(\text{Span}_{\mathbb{C}}(M, a), \mathbb{R}) \quad (27.762)$$

vérifiant  $\|\ell\| = \|f\| = \|f_1\| = \|f_2\|$ .

En utilisant à nouveau le lemme 27.159, nous avons encore une extension

$$\psi(f_2) \in \mathcal{L}_{\mathbb{C}}(\text{Span}_{\mathbb{C}}(M, a), \mathbb{C}) \quad (27.763)$$

vérifiant  $\|\ell\| = \|f\| = \|f_1\| = \|f_2\| = \|\psi(f_2)\|$ .

Et voilà ! La fonctionnelle  $\psi(f_2)$  est celle que nous voulions.  $\square$

### 27.14.5 Prolongement de fonctionnelles (dimension infinie)

Les lemmes 27.169 et 27.170 permettent de prolonger une forme linéaire une dimension réelle ou complexe à la fois. Rien ne nous permet de prolonger d'une infinité de dimensions d'un seul coup. Le théorème de Hahn-Banach va nous permettre de faire une infinité de dimension d'un coup à l'aide du lemme de Zorn.

THOooTZSSooBKfxXE

**Théorème 27.171** (Hahn-Banach[763]).

Soit  $X$ , un espace vectoriel normé sur  $\mathbb{K}$  ( $= \mathbb{R}$  ou  $\mathbb{C}$ ). Soient un sous-espace vectoriel  $M$ , et une fonctionnelle linéaire  $\ell \in \mathcal{L}_{\mathbb{K}}(M, \mathbb{K})$  de norme finie. Il existe  $f \in \mathcal{L}_{\mathbb{K}}(X, \mathbb{K})$  telle que

- (1)  $f(x) = \ell(x)$  pour tout  $x \in M$ ,
- (2)  $\|f\|_{\mathcal{L}(X, \mathbb{K})} = \|\ell\|_{\mathcal{L}(M, \mathbb{K})}$ .

*Démonstration.* Nous considérons l'ensemble  $\mathcal{E}$  des paires  $(N, f)$  telles que

- (1)  $N$  est un sous-espace de  $X$  contenant  $M$ ,
- (2)  $f \in \mathcal{L}(N, \mathbb{K})$
- (3)  $\|f\| = \|\ell\|$
- (4)  $f(x) = \ell(x)$  pour tout  $x \in M$ .

Cet ensemble n'est pas vide parce que  $(M, \ell) \in \mathcal{E}$ . Nous mettons un ordre partiel sur  $\mathcal{E}$  en posant  $(N, f) \leq (N', f')$  si et seulement si  $N \subset N'$  et  $f'|_N = f$ .

- (1)  **$\mathcal{E}$  est inductif** Nous commençons par prouver que  $(\mathcal{E}, \leq)$  est un ensemble inductif<sup>89</sup>. Soit une partie  $\mathcal{F}$  totalement ordonnée dans  $\mathcal{E}$ .

(1a) **L'espace vectoriel**

Nous commençons par poser

$$Y = \bigcup_{(N, f) \in \mathcal{F}} N. \quad (27.764)$$

Et nous prouvons que c'est un espace vectoriel. Soient  $x, y \in Y$ . Supposons  $x \in N_1$  et  $y \in N_2$ . Vu que  $\mathcal{F}$  est totalement ordonné, nous avons  $N_1 \subset N_2$  (ou le contraire). Donc  $x + y \in N_2 \subset Y$ . De même pour  $\lambda x$  avec  $\lambda \in \mathbb{K}$  et  $x \in Y$ .

- (1b) **La fonctionnelle** Nous devons trouver une fonctionnelle  $g$  sur  $Y$ . Soit  $y \in Y$ . Commençons par prouver que l'ensemble

$$\{f(y) \text{ tel que } (N, f) \in \mathcal{F}, y \in N\} \quad (27.765)$$

est un singleton. Soient  $(N_1, f_1), (N_2, f_2) \in \mathcal{F}$  avec  $y \in N_1 \cap N_2$ . Nous supposons que  $(N_1, f_1) \leq (N_2, f_2)$  (sinon c'est le contraire). Alors  $f_2(y) = f_2|_{N_1}(y) = f_1(y)$ . Nous pouvons donc définir

$$\begin{aligned} g: Y &\rightarrow \mathbb{K} \\ y &\mapsto f(y) \end{aligned} \quad (27.766)$$

où  $(N, f) \in \mathcal{F}$  est tel que  $y \in N$ .

---

89. Définition 1.36.



- (1c) **g est linéaire** Soient  $x, y \in Y$ . Nous supposons que  $x \in N_1$  et  $y \in N_2$  avec  $N_1 \subset N_2$ . Donc  $x, y, x + y \in N_2$  et nous avons  $g(x) = f_2(x)$ ,  $g(y) = f_2(y)$  et  $g(x + y) = f_2(x + y)$ . La linéarité de  $f_2$  fait alors le boulot. Même raisonnement pour  $g(\lambda x) = \lambda g(x)$ .
- (1d) **g se restreint à  $\ell$**  Soit  $x \in M$ . Nous avons  $g(x) = f(x)$  pour un couple  $(N, f) \in \mathcal{F}$  vérifiant  $x \in N$ . Vu que  $f$  prolonge  $\ell$  nous avons  $g(x) = f(x) = \ell(x)$ .
- (1e) **Norme de g** Nous devons voir que  $\|g\|_{\mathcal{L}(Y, \mathbb{K})} = \|\ell\|_{\mathcal{L}(M, \mathbb{K})}$ . L'inégalité dans un sens est facile pour qui comprend la norme opérateur<sup>90</sup>. Étant donné que  $M \subset Y$  nous avons

$$\|g\|_{\mathcal{L}(Y, \mathbb{K})} = \sup_{y \in Y} \frac{|g(y)|}{\|y\|} \geq \sup_{y \in M} \frac{|g(y)|}{\|y\|} = \sup_{y \in M} \frac{|\ell(y)|}{\|y\|} = \|\ell\|_{\mathcal{L}(M, \mathbb{K})}. \quad (27.767)$$

L'inégalité dans l'autre sens n'est pas trop compliquée non plus. Prenons  $x \in Y$  vérifiant  $\|x\| = 1$ . Nous considérons  $(N, f) \in \mathcal{E}$  tel que  $x \in N$ . Alors

$$|g(x)| = |f(x)| \leq \|f\| = \|\ell\|. \quad (27.768)$$

Donc  $\|g\| \leq \|\ell\|$ .

- (1f) **Conclusion pour le moment** Nous avons prouvé que  $(Y, g)$  est un majorant de  $\mathcal{F}$ . Donc  $(\mathcal{E}, \leq)$  est un ensemble inductif.

- (2) **Lemme de Zorn** Vu que  $(\mathcal{E}, \leq)$  est un ensemble inductif non vide, il possède un élément maximal par lemme de Zorn 1.37. Nous nommons  $(Y, f)$  un tel élément maximal.
- (3) **Fin de la preuve** Nous devons prouver que  $Y = X$ , de telle sorte que  $f \in \mathcal{L}(Y, \mathbb{K})$  soit définie sur  $X$ . Supposons que  $Y \neq X$ . Dans ce cas nous considérons  $x_1 \in X \setminus Y$ . Suivant que  $\mathbb{K}$  est  $\mathbb{R}$  ou  $\mathbb{C}$ , nous utilisons le lemme 27.169 ou 27.170 pour construire la paire

$$(\text{Span}_{\mathbb{K}}(Y, x_1), f_1) \quad (27.769)$$

qui majore  $(Y, f)$ . Contradiction avec la maximalité de  $(Y, f)$ . Donc  $Y = X$ .

□

PROPOFJPXooWrijbuH

### Proposition 27.172 ([1]).

Soit un espace de Banach  $E$  sur  $\mathbb{K}$  ( $= \mathbb{R}$  ou  $\mathbb{C}$ ). Soit  $a \in E$ . Nous avons

$$\|a\| = \max_{\substack{\varphi \in E' \\ \|\varphi\|=1}} |\varphi(a)|. \quad (27.770)$$

*Démonstration.* Soit  $\varphi \in E'$  tel que  $\|\varphi\| = 1$ . Si  $|\varphi(a)| > \|a\|$  alors

$$\|\varphi\| = \sup_{x \in E} \frac{|\varphi(x)|}{\|x\|} \geq \frac{|\varphi(a)|}{\|a\|} > 1 \quad (27.771)$$

Nous en déduisons que pour tout  $\varphi \in B_{E'}(0, 1)$ ,  $|\varphi(a)| \leq \|a\|$ .

Maintenant nous construisons une application linéaire  $\varphi \in E'$  telle que  $|\varphi(a)| = \|a\|$  et  $\|\varphi\| = 1$ . Pour cela nous considérons le sous-espace  $M = \text{Span}\{a\}$  et l'application

$$\begin{aligned} f: M &\rightarrow \mathbb{K} \\ \lambda a &\mapsto \lambda \|a\|. \end{aligned} \quad (27.772)$$

Cette application vérifie  $f(a) = \|a\|$  et  $\|f\| = 1$  parce que

$$\|f\| = \sup_{x \in M} \frac{|f(x)|}{\|x\|} = \sup_{\lambda \in \mathbb{K}} \frac{|f(\lambda a)|}{\|\lambda a\|} = \frac{|\lambda| \|a\|}{|\lambda| \|a\|} = 1. \quad (27.773)$$

Le théorème de Hahn-Banach 27.171 prolonge  $f$  en un élément  $\varphi \in E'$  vérifiant  $\|\varphi\| = \|f\| = 1$ . Cette fonctionnelle vérifie donc aussi  $\varphi(a) = f(a) = \|a\|$ . □

90. Si ce n'est pas votre cas, vous ne devriez franchement pas être en train de lire ces lignes. Ce n'est que mon avis ; après tout, vous faites comme vous le sentez.

CORooOBDHooJpiBrs

**Corolaire 27.173** ([1]).

Voici trois façons différentes de dire la même chose, par ordre décroissant de frime.

- (1) Le dual d'un espace vectoriel normé sépare les points.
- (2) Les fonctionnelles bornées d'un espace vectoriel normé séparent les points.
- (3) Si  $X$  est un espace vectoriel normé sur  $\mathbb{K}$  ( $= \mathbb{R}$  ou  $\mathbb{C}$ ), et si  $x, y \in X$ , alors il existe une application  $f \in \mathcal{L}(X, \mathbb{K})$  telle que  $f(x) \neq f(y)$ .

*Démonstration.* Le théorème de la base incomplète 4.26 nous permet de considérer une base  $\{e_i\}_{i \in I}$  de  $X$  telle que  $e_0 = x$  et  $e_1 = y$ .

Évacuons quelque objections.

- Si  $x$  et  $y$  sont colinéaires, on complète seulement  $\{x\}$  et ça ne changera rien pour la suite<sup>91</sup>.
- En écrivant «  $e_0$  » et «  $e_1$  » ne prétends pas que  $I$  soit un ensemble de nombres. C'est juste une facilité d'écriture pour éviter de dire « il existe  $\alpha, \beta \in I$  tels que  $x = e_\alpha$  et  $y = e_\beta$  ».

Nous considérons l'application linéaire

$$\begin{aligned} \ell: \text{Span}_{\mathbb{K}}(e_0, e_1) &\rightarrow \mathbb{K} \\ \alpha e_1 + \beta e_2 &\mapsto \alpha. \end{aligned} \quad (27.774)$$

Cette application est parfaitement bornée. Le théorème de Hahn-Banach 27.171 nous permet de considérer une extension  $f$  sur  $X$ . Cette extension a la même norme et est donc bornée (donc continue par la proposition 10.81). C'est donc un élément du dual de  $X$ .

Comme  $f = \ell$  sur  $\text{Span}(e_0, e_1)$ , nous avons  $f(x) \neq f(y)$ . □

## 27.15 Théorème de Tietze

### Définition 27.174.

Si  $E$  et  $F$  sont des espaces normés, une application  $f: E \rightarrow F$  est **presque surjective** si il existe  $\alpha \in ]0, 1[$  et  $C > 0$  tels que pour tout  $y \in \overline{B_F(0, 1)}$ , il existe  $x \in B_E(0, C)$  tel que  $\|y - f(x)\| \leq \alpha$ .

LemBQLooRXhJzK

**Lemme 27.175** ([126]).

Soient  $E$  et  $F$  des espaces de Banach et  $f \in \mathcal{L}^0(E, F)$ <sup>92</sup>. Si  $f$  est presque surjective, alors

ItemTSOooYkxvBui

(1)  $f$  est surjective

ItemTSOooYkxvBui

(2) pour tout  $y \in \overline{B_F(0, 1)}$ , il existe  $x \in \overline{B_E(0, \frac{C}{1-\alpha})}$  tel que  $y = f(x)$ .

Le point (2) est une précision du point (1) : il dit quelle est la taille de la boule de  $E$  nécessaire à obtenir la boule unité dans  $F$ .

*Démonstration.* Soit  $y \in \overline{B_F(0, 1)}$ . Nous allons construire  $x \in B(0, \frac{C}{1-\alpha})$  qui donne  $f(x) = y$ . Ce  $x$  sera la limite d'une série que nous allons construire par récurrence. Pour  $n = 1$  nous utilisons la presque surjectivité pour considérer  $x_1 \in \overline{B_E(0, C)}$  tel que  $\|y - f(x_1)\| \leq \alpha$ . Ensuite nous considérons la récurrence

$$x_n \in \overline{B_E(0, C)} \quad (27.775)$$

tel que

$$\left\| y - \sum_{i=1}^n \alpha^{i-1} f(x_i) \right\| \leq \alpha^n \quad (27.776)$$

Pour montrer que cela existe nous supposons que la série est déjà construite jusqu'à  $n > 1$  :

$$\frac{1}{\alpha^n} \left( y - \sum_{i=1}^n \alpha^{i-1} f(x_i) \right) \in \overline{B_F(0, 1)} \quad (27.777)$$

91. Soyez quand même attentives à ne pas vous laisser enfumer.

92. L'ensemble des applications linéaires continues

À partir de là, par presque surjectivité il existe un  $x_{n+1} \in \overline{B_E(0, C)}$  tel que

$$\left\| \frac{y - \sum_{i=1}^n \alpha^{i-1} f(x_i)}{\alpha^n} - f(x_{n+1}) \right\| \leq \alpha. \quad (27.778)$$

En multipliant par  $\alpha^n$ , le terme  $\alpha^n f(x_{n+1})$  s'intègre bien dans la somme :

$$\left\| y - \sum_{i=1}^{n+1} \alpha^{i-1} f(x_i) \right\| \leq \alpha^{n+1}. \quad (27.779)$$

Nous nous intéressons à une éventuelle limite à la somme des  $\alpha^{n-1}x_n$ . D'abord nous avons la majoration  $\|\alpha^{n-1}x_n\| \leq \alpha^{n-1}C$ , et vu que par la définition de la presque surjectivité  $0 < \alpha < 1$ , la série

$$\sum_{n=1}^{\infty} \alpha^{n-1} x_n \quad (27.780)$$

converge absolument<sup>93</sup> parce que la suite des normes est une suite géométrique de raison  $\alpha$ . Vu que  $E$  est de Banach, la convergence absolue implique la convergence simple (la suite des sommes partielles est de Cauchy et Banach est complet). Nous posons

$$x = \sum_{n=1}^{\infty} \alpha^{n-1} x_n \in E, \quad (27.781)$$

et en termes de normes, ça vérifie

$$\|x\| \leq \sum_{n=1}^{\infty} \alpha^{n-1} \|x_n\| \leq C \sum_{n=1}^{\infty} \alpha^{n-1} = \frac{C}{1-\alpha}. \quad (27.782)$$

Donc c'est bon pour avoir  $x \in B(0, \frac{C}{1-\alpha})$ . Nous devons encore vérifier que  $y = f(x)$ . Pour cela nous remarquons que

$$\left\| y - f\left(\sum_{n=1}^N \alpha^{n-1} x_n\right) \right\| \leq \alpha^N. \quad (27.783)$$

Nous pouvons prendre la limite  $N \rightarrow \infty$  et permuter  $f$  avec la limite (par continuité de  $f$ ). Vu que  $0 < \alpha < 1$  nous avons

$$\|y - f(x)\| = 0. \quad (27.784)$$

□

ThoFFQooGvcLzJ

**Théorème 27.176** (Tietze[126, 555, 764]).

Soit un espace métrique  $(X, d)$  et un fermé  $Y \subset X$ . Soit  $g_0 \in C^0(Y, \mathbb{R})$ . Alors  $g_0$  admet un prolongement continu sur  $X$ .

*Démonstration.* Soit l'opération de restriction

$$\begin{aligned} T: (C_b^0(X, \mathbb{R}), \|\cdot\|_{\infty}) &\rightarrow (C_b^0(Y, \mathbb{R}), \|\cdot\|_{\infty}) \\ f &\mapsto f|_Y. \end{aligned} \quad (27.785)$$

L'application  $T$  est évidemment linéaire. Elle est de plus bornée pour la norme opérateur usuelle donnée par la proposition 10.63 parce que  $\|T(f)\| \leq \|f\| < \infty$ . L'application  $T$  est alors continue par la proposition 10.81.

(1) **Presque surjection**

Soit  $g \in C_b^0(Y, \mathbb{R})$  avec  $\|g\|_{\infty} \leq 1$ . Nous posons

$$Y^+ = \{x \in Y \text{ tel que } \frac{1}{3} \leq g(x) \leq 1\} \quad (27.786a)$$

$$Y^- = \{x \in Y \text{ tel que } -1 \leq g(x) \leq -\frac{1}{3}\}. \quad (27.786b)$$

---

93. Définition 10.109.

Nous considérons alors

$$\begin{aligned} f: X &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{3} \frac{d(x, Y^-) - d(x, Y^+)}{d(x, Y^-) + d(x, Y^+)} \end{aligned} \quad (27.787)$$

Vu qu'en valeur absolue le dénominateur est plus grand que le numérateur nous avons  $\|f\|_\infty \leq \frac{1}{3}$ . Notons que

- Si  $x \in Y^+$  alors  $f(x) = \frac{1}{3}$  et  $g(x) \in [\frac{1}{3}, 1]$ ;
- Si  $x \in Y^-$  alors  $f(x) = -\frac{1}{3}$  et  $g(x) \in [-1, -\frac{1}{3}]$ ;
- Si  $x$  n'est ni dans  $Y^+$  ni dans  $Y^-$  alors nous avons<sup>94</sup>  $g(x) \in [-\frac{1}{3}, \frac{1}{3}]$  et donc  $|f(x) - g(x)| \leq |f(x)| + |g(x)| \leq \frac{2}{3}$ .

Dans les trois cas nous avons  $|f(x) - g(x)| \in [0, \frac{2}{3}]$  pour tout  $x \in X$ . Cela prouve que

$$\|T(f) - g\|_{Y, \infty} \leq \frac{2}{3}. \quad (27.788)$$

En résumé nous avons pris  $g$  dans la boule  $\overline{B(0, 1)}$  de  $(C_b^0(Y, \mathbb{R}), \|\cdot\|_\infty)$  et nous avons construit une fonction  $f$  dans la boule  $\overline{B(0, \frac{1}{3})}$  de  $(C_b^0(X, \mathbb{R}), \|\cdot\|_\infty)$  telle que  $\|T(f) - g\|_\infty \leq \frac{2}{3}$ . L'application  $T$  est donc une presque surjection avec  $\alpha = \frac{2}{3}$  et  $C = \frac{1}{3}$ .

- (2) **Extension avec  $\|\cdot\|_\infty \leq 1$**  SPITEMooSTXMooldockk Nous montrons que si  $g \in C^0(Y, \mathbb{R})$  vérifie  $\|g\|_\infty \leq 1$ , alors il existe une extension  $f \in C^0(X, \mathbb{R})$  vérifiant  $\|f\|_\infty \leq 1$ .

La proposition 12.383 nous assure que les espaces  $C_b^0(X, \mathbb{R})$  et  $C_b^0(Y, \mathbb{R})$  sont de Banach (complets), et le lemme 27.175 nous dit alors que  $T$  est surjective et que pour tout  $g \in \overline{B(0, 1)}$ , il existe

$$f \in \overline{B\left(0, \frac{1/3}{1 - \frac{2}{3}}\right)} = \overline{B(0, 1)}. \quad (27.789)$$

telle que  $g = T(f)$ .

Pour le redire avec des mots plus simples, nous avons prouvé que si  $g: Y \rightarrow \mathbb{R}$  vérifie  $\|g\|_\infty \leq 1$ , il existe une extension  $f: X \rightarrow \mathbb{R}$  telle que  $\|f\|_\infty \leq 1$ .

- (3) **Extension avec  $|\cdot| < 1$**  SPITEMooZBGTooKbyuch Nous montrons que si  $g \in C^0(Y, \mathbb{R})$  vérifie  $|g| < 1$ , alors il existe une extension  $f \in C^0(X, \mathbb{R})$  vérifiant  $|f| < 1$ .

Soit  $g \in B_{C_b^0(Y)}(0, 1)$  et son prolongement  $h \in \overline{B_{C_b^0(X)}(0, 1)}$  du point (2). Si  $|h| < 1$  alors nous avons le résultat en posant  $f = g$ .

Si au contraire il existe  $z \in X$  tel que  $|h(z)| = 1$ , nous posons

$$Z = \{x \in X \text{ tel que } |h(x)| = 1\}. \quad (27.790) \quad \text{EQooIVXAooZQZole}$$

Nous avons  $Y \cap Z = \emptyset$  parce que nous avons  $h = g$  sur  $Y$  et nous avons choisi  $\|g\|_\infty < 1$ . Par ailleurs  $Y$  est fermé par hypothèse et  $Z$  est fermé parce que  $h$  est continue; par conséquent  $Y \cap Z$  est fermé, donc<sup>95</sup>

$$\bar{Y} \cap \bar{Z} = Y \cap Z = \emptyset. \quad (27.791)$$

Nous posons

$$\begin{aligned} u: X &\rightarrow \mathbb{R}^+ \\ x &\mapsto \frac{d(x, Z)}{d(x, Y) + d(x, Z)} \end{aligned} \quad (27.792)$$

Le dénominateur n'est pas nul parce qu'il faudrait  $d(x, Y) = d(x, Z) = 0$ , ce qui demanderait  $x \in \bar{Y} \cap \bar{Z}$ , ce qui n'est pas possible. Nous posons  $f = uh$ . Si  $x \in Y$  alors  $u(x) = 1$ , donc  $f$  est encore un prolongement de  $g$ . De plus  $f$  est encore continue, et donc encore un bon candidat. Enfin si  $x$  est hors de  $Y$  alors  $d(x, Y) > 0$  (strictement parce que  $Y$  est fermé) et donc  $0 < u(x) < 1$ , ce qui donne  $|f(x)| < |h(x)| \leq 1$ .

Et voilà.

94. Nous rappelons que  $\|g\| = 1$ , donc  $g(x)$  est forcément entre  $-1$  et  $1$ .

95. Si vous avez l'intention de dire que  $\bar{Y} \cap \bar{Z} = \bar{Y} \cap \bar{Z} = Y \cap Z = \emptyset$ , allez d'abord voir l'exemple 7.177. Ici c'est correct parce que  $Y$  et  $Z$  sont fermés.

(4) **Démonstration du théorème**

Nous sommes maintenant prêts à prouver le théorème. Soit  $g_0 \in C^0(Y, \mathbb{R})$ . Nous allons nous ramener au cas de la boule unité ouverte en utilisant un homéomorphisme<sup>96</sup>  $\phi: \mathbb{R} \rightarrow ]-1, 1[$ . L'application  $g = \phi \circ g_0$  vérifie  $|g(y)| < 1$  pour tout  $y \in Y$ . Par le point (3), il existe une extension  $f \in C^0(X, \mathbb{R})$  vérifiant  $|f(x)| < 1$  pour tout  $x \in X$ . Attention : cette application  $f$  prolonge  $g$  et non  $g_0$ .

Nous posons finalement  $f_0 = \phi^{-1} \circ f$ . Cela est encore une application dans  $C^0(X, \mathbb{R})$ , mais cette fois elle prolonge  $g_0$ . En effet si  $y \in Y$ , en tenant compte de  $f(y) = g(y)$ , nous avons

$$f_0(y) = (\phi^{-1} \circ f)(y) = \phi^{-1}(f(y)) = \phi^{-1}(g(y)) = \phi^{-1}((\phi \circ g_0)(y)) = g_0(y). \quad (27.793)$$

□

**27.15.1 Pas de bicontinues entre dimensions différentes**

THOooLGJMooIYzOBD

**Théorème 27.177 ([765]).**

Soit une application continue et injective  $f: \overline{B(0, 1)} \rightarrow \mathbb{R}^n$ . Alors

$$f(0) \in \text{Int} \left( f(\overline{B(0, 1)}) \right). \quad (27.794)$$

*Démonstration.* Soit une application continue et injective  $f_0: \overline{B(0, 1)} \rightarrow \mathbb{R}^n$ . Soit  $\epsilon < 0$ . Si nous supposons que  $f_0(0)$  n'est pas à l'intérieur de  $f_0(\overline{B(0, 1)})$ , alors il existe un  $c \in \mathbb{R}^n$  hors de  $f_0(\overline{B(0, 1)})$  et tel que  $\|c - f_0(0)\| < \epsilon$ .

Nous posons maintenant  $f(x) = f_0(x) - c$ . Nous avons

$$\begin{cases} \|f(0)\| < \epsilon \\ 0 \notin f(\overline{B(0, 1)}). \end{cases} \quad \begin{matrix} (27.795a) \\ (27.795b) \end{matrix}$$

Vu que  $f$  est continue et que  $\overline{B(0, 1)}$  est compact, la partie  $f(\overline{B(0, 1)})$  est compacte (théorème 7.274). L'application restreinte  $f: \overline{B(0, 1)} \rightarrow f(\overline{B(0, 1)})$  est une bijection continue entre compacts. Son inverse est alors également continue (lemme 7.277).

Mais comme  $f(\overline{B(0, 1)})$  est un fermé de  $\mathbb{R}^n$ , le théorème de Tietze (27.176) prolonge son inverse en une application continue

$$G: \mathbb{R}^n \rightarrow \mathbb{R}^n. \quad (27.796)$$

Nous avons  $G(f(0)) = 0$ . Nous choisissons  $\epsilon$  assez petit pour que  $\|G(y)\| < 0.1$  dès que  $\|y - f(0)\| \leq 2\epsilon$ . Comme  $\|f(0)\| < \epsilon$ , nous avons aussi

$$\|y\| \leq \epsilon \Rightarrow \|y - f(0)\| \leq 2\epsilon \Rightarrow \|G(y)\| < 0.1. \quad (27.797)$$

Autrement dit,

$$G(\overline{B(0, \epsilon)}) \subset B(0, \frac{1}{10}). \quad (27.798)$$

Nous posons

$$\Sigma_1 = \{y \in f(\overline{B(0, 1)}) \text{ tel que } \|y\| \geq \epsilon\}, \quad (27.799)$$

ainsi que

$$\Sigma_2 = \{y \in \mathbb{R}^n \text{ tel que } \|y\| = \epsilon\}, \quad (27.800)$$

et  $\Sigma = \Sigma_1 \cup \Sigma_2$ .

- (1) La partie  $\Sigma_2$  est compacte en tant que fermée et bornée<sup>97</sup>. La partie  $\Sigma_1$  est compacte en tant que intersection entre le compact  $f(\overline{B(0, 1)})$  et le fermé  $\|y\| \geq \epsilon$  (corolaire 7.122). Finalement la partie  $\Sigma$  est également compacte<sup>98</sup>.

96. Existence dans le lemme 18.41.

97. Borel-Lebesgue 9.24.

98. Union de compacts, lemme 7.124.

- (2) La partie  $\Sigma$  ne contient pas  $f(0)$  parce que  $\|f(0)\| < \epsilon$  tandis que les éléments  $y \in \Sigma$  vérifient  $\|y\| \geq \epsilon$ .
- (3) La partie  $\Sigma_1$  ne contient pas 0. En effet, vu que  $G$  étend  $f^{-1}$ , si  $x \in G(\Sigma_1)$ , nous avons  $f(x) \in \Sigma_1$  et en particulier  $\|f(x)\| \geq \epsilon$ . Vu que  $\|f(0)\| < \epsilon$ , nous avons  $x \neq 0$ .
- (4) La partie  $\Sigma_1$  est compacte,  $G$  est continue dessus et ne prend pas la valeur 0. Il existe donc  $\delta > 0$  tel que

$$G(\Sigma_1) > \delta. \quad (27.801)$$

Nous considérons un tel  $\delta$  qui vérifie de plus  $\delta < 1/10$ .

Nous introduisons l'application

$$\begin{aligned} \phi: f(\overline{B(0,1)}) &\rightarrow \Sigma \\ y &\mapsto \max(\epsilon/\|y\|, 1)y. \end{aligned} \quad (27.802)$$

Prouvons que  $\phi$  prend effectivement ses valeurs dans  $\Sigma$ . Il y a deux possibilités : soit  $\|y\| < \epsilon$ , soit  $\|y\| \geq \epsilon$ . Si  $\|y\| < \epsilon$ , alors  $\phi(y) = \frac{\epsilon}{\|y\|}y$ , ce qui donne  $\|\phi(y)\| = \epsilon$ , et donc  $\phi(y) \in \Sigma_2$ . Si  $\|y\| \geq \epsilon$ , alors le max est 1 et nous avons  $\phi(y) = y$ . Dans ce cas, nous avons d'une part  $\|\phi(y)\| = \|y\| \geq \epsilon$  et d'autre part  $\phi(y) = y \in f(\overline{B(0,1)})$  ; donc  $\phi(y) \in \Sigma_1$ .

Le théorème de Stone-Weierstrass 12.447 nous permet de considérer un polynôme (à  $n$  variables)  $P: \mathbb{R}^n \rightarrow \mathbb{R}^n$  vérifiant

$$\|P - G\|_{\Sigma} < \delta. \quad (27.803)$$

Un tel polynôme ne s'annule pas sur  $\Sigma_1$  parce que  $\|G\|_{\Sigma_1} > \delta$ .

Le polynôme  $P$  est de classe  $C^\infty$  (proposition 10.288). Vu que  $\Sigma_2$  est de mesure nulle, la proposition 14.280 dit que  $P(\Sigma_2)$  est de mesure nulle dans  $\mathbb{R}^n$ .

À l'aide de la proposition 14.281, nous choisissons  $s \in \mathbb{R}^n$  tel que  $P_2(x) = P(x+s)$  ne s'annule ni sur  $\Sigma_1$  ni sur  $\Sigma_2$ . Vu que  $P$  est uniformément continue sur  $\Sigma$  (théorème de Heine 12.80), il existe  $M$  (majorant de la norme de la dérivée sur  $\Sigma$ ) tel que

$$\|P(x) - P_2(x)\| = \|P(x+s) - P(x)\|_{\Sigma} \leq sM. \quad (27.804)$$

En choisissant  $s < \epsilon'/M$ , nous avons

$$\|P - P_2\|_{\Sigma} < \epsilon'. \quad (27.805)$$

En prenant  $\epsilon'$  assez petit<sup>99</sup>, nous avons encore

$$\|P_2 - G\|_{\Sigma} < \delta. \quad (27.806)$$

Nous introduisons maintenant

$$\begin{aligned} \tilde{G}: f(\overline{B(0,1)}) &\rightarrow \mathbb{R}^n \\ \tilde{G} &= P_2 \circ \phi. \end{aligned} \quad (27.807)$$

Nous allons prouver que  $\|G(y) - \tilde{G}(y)\| \leq 0.3$  pour tout  $y \in f(\overline{B(0,1)})$ .

- (1) Si  $\|y\| > \epsilon$

Si  $\|y\| > \epsilon$  et  $y \in f(\overline{B(0,1)})$ , alors

$$\tilde{G}(y) = (P_2 \circ \phi)(y) = P(\min(\epsilon/\|y\|, 1)y). \quad (27.808)$$

Vu que  $\|y\| > \epsilon$ , le max est 1 et nous avons  $\tilde{G}(y) = P_2(y)$ , et donc, en utilisant (27.806),

$$\|G(y) - \tilde{G}(y)\| = \|G(y) - P_2(y)\| < \delta. \quad (27.809)$$

Quitte à choisir  $\delta$  plus petit, nous pouvons supposer  $\delta < 0.3$ .

99. Plus petit que la différence entre  $\delta$  et  $\|P - G\|_{\Sigma}$ .

(2) Si  $\|y\| \leq \epsilon$  Nous considérons maintenant le cas  $\|y\| \leq \epsilon$  et  $y \in \overline{b(B(0,1))}$ .

Montrons que  $\|G(y) - \tilde{G}(y)\| < \delta$  pour tout  $y \in \overline{B(0,1)}$  tel que  $\|y\| > \epsilon$ . Nous avons

$$\|G(y) - \tilde{G}(y)\| = \|G(y) - P_2(\phi(y))\| \quad (27.810a)$$

$$\leq \|G(y) - G(\phi(y))\| + \|G(\phi(y)) - P_2(\phi(y))\| \quad (27.810b)$$

$$\leq \|G(y)\| + \|G(\phi(y))\| + \|(G - P_2)(\phi(y))\|. \quad (27.810c)$$

Vu que  $\|y\| \leq \epsilon$ , nous avons  $\phi(y) = \frac{\epsilon}{\|y\|}y$ , et donc  $\|\phi(y)\| = \epsilon$ . De ce fait,  $\phi(y) \in \Sigma_2$  et nous avons  $\|(G - P_2)(\phi(y))\| \leq \delta$ .

Si nous choisissons  $\delta < 0.1$  nous avons donc encore

$$\|G(y) - \tilde{G}(y)\| \leq 2 \times 0.1 + \delta \leq 0.3. \quad (27.811)$$

Bref, pour tout  $y \in \overline{f(B(0,1))}$  nous avons

$$\|G(y) - \tilde{G}(y)\| \leq 0.3. \quad (27.812)$$

Ça va être le moment de demander au théorème de Brouwer d'intervenir.

(1) **Définition de  $g$**

Vu que  $\tilde{G}$  est définie sur  $\overline{f(B(0,1))}$ , nous pouvons considérer l'application

$$\begin{aligned} g: \overline{B(0,1)} &\rightarrow \mathbb{R}^n \\ x &\mapsto x - \tilde{g}(f(x)). \end{aligned} \quad (27.813)$$

Notons pour  $x \in \overline{B(0,1)}$ , nous avons  $x = G(f(x))$ , et donc nous pouvons aussi écrire

$$g(x) = (G \circ f)(x) - (\tilde{G} \circ f)(x). \quad (27.814)$$

(2)  $g(x) \in \overline{B(0,1)}$

Nous avons déjà vu en (27.812) que  $\|G - \tilde{G}\|_{f(\overline{B(0,1)})} < 0.3 < 1$ . L'équation (27.814) dit alors que  $g(x) \in \overline{B(0,1)}$ .

(3) **Brouwer** Nous pouvons appliquer le théorème de Brouwer (20.32) à la fonction  $g$ . Il existe  $x_0 \in \overline{B(0,1)}$  tel que  $g(x_0) = x_0$ , c'est-à-dire  $x_0 - \tilde{G}(f(x_0)) = x_0$ . Nous en déduisons que

$$\tilde{G}(f(x_0)) = 0. \quad (27.815)$$

(4) **La contradiction** Rappel :  $\tilde{G} = P_2 \circ \phi$ . Si  $y \in \overline{B(0,1)}$ , alors  $\phi(y) \in \Sigma$ , et nous avons tout fait pour que  $P_2$  ne s'annule pas sur  $\Sigma$ .

Et voilà la contradiction.

□

### Remarque 27.178.

Il n'existe pas de bijection bicontinues d'un ouvert de  $\mathbb{R}^m$  vers un ouvert de  $\mathbb{R}^n$  si  $m \neq n$ . Il n'y a donc pas de notion de difféomorphismes entre ouverts de dimensions différentes.

Je crois que c'est une conséquence facile de 27.177. Écrivez-moi si vous pouvez rédiger une preuve.

## 27.16 Dualité, réflexivité et théorème de représentation de Riesz

Dans la suite  $E'$  est le dual topologique, c'est-à-dire l'espace des formes linéaires et continues sur  $E$ . Nous notons également  $V''$  le dual de  $(V', \|\cdot\|)$ . Certes en tant qu'ensembles,  $(V', *)$  et  $(V', \|\cdot\|)$  sont identiques, mais comme ils n'ont pas la même topologie, les duals ne sont pas les mêmes.

Bref,  $V''$  est l'ensemble des applications linéaires continues  $(V', \|\cdot\|) \rightarrow \mathbb{C}$ . Et lorsque nous disons  $\mathbb{C}$  ici, ça peut aussi bien être  $\mathbb{R}$  selon le contexte.

De plus nous considérons que  $V''$  la norme opérateur qui dérive de la norme de  $V'$ , laquelle dérive de la norme vectorielle sur  $V$ .

PROPooMAQSooCGFBBM

**Proposition-Définition 27.179** (Espace réflexif).

Soit un espace vectoriel normé  $V$  sur  $\mathbb{R}$  ou  $\mathbb{C}$ . Nous considérons l'application

$$\begin{aligned} J: V &\rightarrow V'' \\ J(x)\varphi &= \varphi(x). \end{aligned} \tag{27.816}$$

EQooRQIMooZdVkdQ

(1) L'application  $J$  est bien définie :  $J(x): V' \rightarrow \mathbb{C}$  est continue.

ITEMooNVVSooNFXgnE

(2) L'application  $J$  est continue.

ITEMooKURHooZZWpbu

(3) Elle est injective.

ITEMooTFYVooKhMOjp

Lorsque  $J$  est bijective, l'espace  $V$  est dit **réflexif**.

*Démonstration.* Point par point.

(1) **(1)** Nous commençons par montrer que  $J(x): (V', \|\cdot\|) \rightarrow \mathbb{C}$  est continue pour chaque  $x \in V$ .

Soit une suite  $\varphi_k \xrightarrow{\|\cdot\|} 0$ . Nous avons :

$$J(x)\varphi_k = \varphi_k(x) \leq \|\varphi_k\| \|x\| \rightarrow 0 \tag{27.817}$$

où vous aurez noté l'utilisation du lemme 10.77. Cela prouve que  $J(x)$  est continue et donc que  $J$  est bien à valeurs dans  $V''$ .

(2) **(2)**

Soit une suite  $x_k \xrightarrow{V} 0$ , et étudions  $\|J(x_k)\|$  pour la norme dans  $V''$ . Nous posons  $x'_k = x_k / \|x_k\|$  et nous calculons (encore une fois, nous écrivons «  $\mathbb{C}$  », mais ça pourrait être  $\mathbb{R}$ )

$$\|J(x_k)\| = \sup_{\|\varphi\|=1} |J(x_k)\varphi|_{\mathbb{C}} = \sup_{\|\varphi\|=1} |\varphi(x_k)| = \|x_k\| \sup_{\|\varphi\|=1} |\varphi(x'_k)| \leq \|x_k\| \rightarrow 0. \tag{27.818}$$

La dernière inégalité pourrait être sans doute une égalité<sup>100</sup>, mais nous n'en avons pas besoin ici.

(3) **(3)** Soient  $x \neq y$  dans  $V$ . Le corolaire 27.173 nous permet de considérer un élément  $\varphi \in V'$  tel que  $\varphi(x) \neq \varphi(y)$ . Nous avons alors

$$J(x)\varphi \neq J(y)\varphi, \tag{27.819}$$

et donc  $J(x) \neq J(y)$ . Cela prouve que  $J$  est injective.

□

PROPooQGHTooLMFCKt

**Proposition 27.180** ([1, 237]).

Soit un espace de Banach  $E$ .

(1) L'application  $J: E \rightarrow E''$  est une isométrie.

ITEMooFTFBooMjahZi

(2) Si  $E$  est réflexif,

$$J(\overline{B_E(0, 1)}) = \overline{B_{E''}(0, 1)}. \tag{27.820}$$

100. Écrivez-moi si vous en êtes certain.



*Démonstration.* Point par point.

- (1) **Isométrie** Soit  $x \in E$ . En utilisant la définition de la norme opérateur et la proposition 27.172 nous avons :

$$\|J(x)\|_{E''} = \sup_{\substack{\varphi \in E' \\ \|\varphi\|=1}} |J(x)\varphi| = \sup_{\substack{\varphi \in E' \\ \|\varphi\|=1}} |\varphi(x)| = \|x\|. \quad (27.821)$$

Donc  $J$  est une isométrie.

- (2) **Pour les boules** Deux inclusions

(2a)  $J(\overline{B_E(0,1)}) \subset \overline{B_{E''}(0,1)}$  Soit  $x \in \overline{B_E(0,1)}$ . Nous savons déjà que  $J$  est isométrique. Donc  $\|J(x)\|_{E''} = \|x\|_E \leq 1$ .

(2b)  $J(\overline{B_{E''}(0,1)}) \subset \overline{B_E(0,1)}$  Soit  $\xi \in \overline{B_{E''}(0,1)}$ . Par hypothèse  $E$  est réflexif ; l'application  $J: E \rightarrow E''$  est donc bijective et nous pouvons considérer  $x = J^{-1}(\xi)$ . En utilisant encore le fait que  $J$  est isométrique,

$$\|x\|_E = \|J(x)\|_{E''} = \|\xi\|_{E''} \leq 1. \quad (27.822)$$

Donc  $J^{-1}(\xi) = x \in \overline{B_E(0,1)}$ .

□

PROP00HOKQooDWpTqM

**Proposition 27.181** ([220]).

Si  $E$  est un espace de Banach, alors  $J(E)$  est fermé dans  $E''$ .

*Démonstration.* Soit une suite convergente  $\xi_n \xrightarrow{E''} \xi$  dans  $J(E)$ . Vu que  $J$  est injective et surjective sur  $J(E)$ , pour chaque  $n$  il existe un unique  $x_n$  tel que  $\xi_n = J(x_n)$ .

Vu que  $(\xi_n)$  est une suite convergente dans l'espace normé  $E''$ , c'est une suite de Cauchy par la proposition 7.322. Étant donné que  $J$  est isométrique, la suite  $(x_n)$  est également de Cauchy. Et donc  $(x_n)$  converge dans  $E$  parce que  $E$  est complet (il est de Banach par hypothèse). Il existe donc  $x \in E$  tel que  $x_n \xrightarrow{E} x$ .

Étant donné que  $J$  est continue, elle commute avec la limite :

$$\xi = \lim_n \xi_n = \lim_n J(x_n) = J(\lim_n x_n) = J(x). \quad (27.823)$$

Donc  $\xi = J(x) \in J(E)$ . Cela prouve que  $J(E)$  est fermé par la proposition 7.80.

□

PROP00JJINooXeUcQP

**Proposition 27.182** ([1]).

L'application

$$J: (E, \sigma(E, E')) \rightarrow (E'', \sigma(E'', E')) \quad (27.824)$$

est continue.

*Démonstration.* Nous utilisons la proposition 7.61 en se souvenant que la topologie  $\sigma(E'', E')$  est la topologie initiale des évaluations

$$\begin{aligned} ev_\varphi: E'' &\rightarrow \mathbb{K} \\ \xi &\mapsto \xi(\varphi). \end{aligned} \quad (27.825)$$

Autrement dit, toutes ces applications sont  $\sigma(E'', E')$ -continues. La proposition 7.61 dit que  $J$  sera continue si toutes les applications  $ev_\varphi \circ J: (E, \sigma(E, E')) \rightarrow \mathbb{K}$  sont continues. Nous avons

$$(ev_\varphi \circ J)(x) = J(x)(\varphi) = \varphi(x). \quad (27.826)$$

Autrement dit,  $ev_\varphi \circ J = \varphi$ . Et  $\varphi$  est  $\sigma(E, E')$ -continue. Donc on est bon.

□

LEMooWEMFooEHlaxY

**Lemme 27.183** ([1]).

Soient des espaces de Banach  $V$  et  $W$ . Si  $\alpha: V \rightarrow W$  est une bijection linéaire et isométrique, alors l'application définie par

$$\begin{aligned}\beta: V' &\rightarrow W' \\ \beta(\varphi)y &= \varphi(\alpha^{-1}(y))\end{aligned}\tag{27.827}$$

EQooTSVHooEQGhNw

pour tout  $\varphi \in V'$  et  $y \in W$  est une bijection linéaire isométrique.

*Démonstration.* En plusieurs points.

- (1)  **$\alpha^{-1}$  est continue** Le théorème d'isomorphisme de Banach 27.1 implique que  $\alpha^{-1}$  est continue parce que  $\alpha$  est bijective, continue et linéaire entre deux espaces de Banach.
- (2)  **$\beta$  prend ses valeurs dans  $W'$**  Nous devons prouver que, si  $\varphi \in V'$ , alors  $\beta(\varphi)$  est bien un élément de  $W'$ . Autrement dit, nous devons prouver que  $\beta(\varphi): W \rightarrow \mathbb{C}$  est continue. Considérons une suite  $y_k \xrightarrow{W} 0$ . Alors, vu que  $\alpha$  est continue, nous avons aussi  $\alpha^{-1}(y_k) \xrightarrow{V} 0$ , et donc

$$\varphi(\alpha^{-1}(y_k)) \xrightarrow{\mathbb{C}} 0 \tag{27.828}$$

parce que  $\varphi$  est continue. Donc  $\beta(\varphi)$  est continue.

- (3)  **$\beta$  est linéaire** Parce que  $\alpha$  l'est.
- (4)  **$\beta$  est isométrique** Nous devons prouver que la norme de l'application linéaire  $\beta(\varphi): W \rightarrow \mathbb{C}$  est la même que celle de  $\varphi$ . Vu que  $\alpha: V \rightarrow W$  est une bijection isométrique, nous avons

$$\{\alpha^{-1}(y) \text{ tel que } \|y\| = 1, y \in W\} = \{x \in V \text{ tel que } \|x\| = 1\}. \tag{27.829}$$

Donc nous pouvons faire le calcul suivant :

$$\|\beta(\varphi)\| = \sup_{\substack{y \in W \\ \|y\|=1}} \|\beta(\varphi)y\| = \sup_{\substack{y \in W \\ \|y\|=1}} |\varphi(\alpha^{-1}(y))| = \sup_{\substack{x \in V \\ \|x\|=1}} |\varphi(x)| = \|\varphi\|. \tag{27.830}$$

□

La proposition suivante dit que la notion d'être réflexif passe aux isomorphismes d'espaces vectoriels normés.

PROPooVRQKooLdmajh

**Proposition 27.184** ([1]).

Soient des espaces vectoriels normés  $V$  et  $W$  ainsi qu'une bijection linéaire isométrique  $\alpha: V \rightarrow W$ . L'espace  $W$  est réflexif<sup>101</sup> si et seulement si  $V$  l'est.

*Démonstration.* Nous allons prouver que si  $W$  est réflexif, alors  $V$  est réflexif. Pour l'implication inverse, il suffira de noter que  $\alpha^{-1}: W \rightarrow V$  est une bijection linéaire isométrique.

En plusieurs points.

- (1) **Quelques applications** Nous notons  $J_V$  l'application de la proposition 27.179 :

$$\begin{aligned}J_V: V &\rightarrow V'' \\ J_V(v)\varphi &= \varphi(v),\end{aligned}\tag{27.831}$$

et  $J_W$  l'application correspondante pour  $W$ . Notre hypothèse est que  $J_W$  est bijective.

Nous posons

$$\begin{aligned}\beta: V' &\rightarrow W' \\ \beta(\varphi)w &= \varphi(\alpha^{-1}(w)),\end{aligned}\tag{27.832}$$

et

$$\begin{aligned}\gamma: V'' &\rightarrow W'' \\ \gamma(\sigma)\varphi &= \sigma(\beta^{-1}(\varphi)).\end{aligned}\tag{27.833}$$

L'application  $\beta$  est une bijection linéaire isométrique par le lemme 27.183 ; l'application  $\gamma$  l'est aussi, par le même lemme appliqué à  $\beta: V' \rightarrow W'$ .

101. Définition 27.179.

(2) **Le  $S$**  Nous posons

$$\begin{aligned} S: V &\rightarrow V'' \\ S(v)\varphi &= J_W(\alpha(v))\beta(\varphi). \end{aligned} \quad (27.834)$$

Nous allons montrer que  $S = J_V$  et que  $S$  est bijectif.

(3)  **$S = J_V$**  Cela est un calcul :

$$J_W(\alpha(v))\beta(\varphi) = \beta(\varphi)(\alpha(v)) \quad (27.835a)$$

$$= \varphi(\alpha^{-1}(\alpha(v))) \quad (27.835b)$$

$$= \varphi(v) \quad (27.835c)$$

$$= J_V(v)\varphi. \quad (27.835d)$$

(4)  **$S$  est injective** L'application  $J_V$  est toujours injective par la proposition 27.179.

(5)  **$S$  est surjective** Soit  $\sigma \in V''$ . Nous devons prouver l'existence de  $v \in V$  tel que  $S(v) = \sigma$ .  
Vu que  $J_W: W \rightarrow W''$  est surjective et que  $\gamma(\sigma) \in W''$ , il existe  $w \in W$  tel que

$$J_W(w) = \gamma(\sigma). \quad (27.836)$$

Un simple calcul montre alors que  $S(\alpha^{-1}(w)) = \sigma$  :

$$S(\alpha^{-1}(w))\varphi = J_W(w)\beta(\varphi) \quad (27.837a)$$

$$= \gamma(\sigma)(\beta(\varphi)) \quad (27.837b)$$

$$= \sigma(\beta^{-1}(\beta(\varphi))) \quad (27.837c)$$

$$= \sigma(\varphi). \quad (27.837d)$$

Nous avons donc bien que  $S(v) = \sigma$  en posant  $v = \alpha^{-1}(w)$ .

(6) **Conclusion** L'application  $S = J_V$  est surjective ; l'espace  $V$  est donc réflexif par la proposition 27.179. □

Voici déjà un bel énoncé. Pour des espaces mesurés  $(\Omega, \mathcal{A}, \mu)$  plus généraux, voir l'arme totale en le théorème 27.187.

PropOAVooYZSodR

**Proposition 27.185** ([432], thème 21).

Soit  $1 < p < 2$  et  $q$  tel que  $\frac{1}{p} + \frac{1}{q} = 1$ . L'application

$$\begin{aligned} \Phi: L^q([0, 1]) &\rightarrow L^p([0, 1])' \\ \Phi_g(f) &= \int_{[0, 1]} f \bar{g}. \end{aligned} \quad (27.838)$$

est une isométrie linéaire surjective.

*Démonstration.* Pour la simplicité des notations nous allons noter  $L^2$  pour  $L^2([0, 1])$ , et pareillement pour  $L^p$ .

(1)  **$\Phi_g$  est un élément de  $(L^p)'$**

Si  $f \in L^p$  et  $g \in L^q$  nous devons prouver que  $\Phi_g(f)$  est bien définie. Pour cela nous utilisons l'inégalité de Hölder<sup>102</sup> qui dit que  $fg \in L^1$  ; par conséquent la fonction  $f\bar{g}$  est également dans  $L^1$  et nous avons

$$|\Phi_g(f)| \leq \int_{[0, 1]} |f\bar{g}| = \|fg\|_1 \leq \|f\|_p \|g\|_q. \quad (27.839)$$

---

102. Proposition 27.39.

En ce qui concerne la norme de l'application  $\Phi_g$  nous avons tout de suite

$$\|\Phi_g\| = \sup_{\|f\|_p=1} |\Phi_g(f)| \leq \|g\|_q. \quad (27.840)$$

Cela signifie que l'application  $\Phi_g$  est bornée et donc continue par la proposition 10.81. Nous avons donc bien  $\Phi_g \in (L^p)'$ .

(2) **Isométrie**

Afin de prouver que  $\|\Phi_g\| = \|g\|_q$  nous allons trouver une fonction  $f \in L^p$  telle que  $\frac{|\Phi_g(f)|}{\|f\|_p} = \|g\|_q$ . De cette façon nous aurons prouvé que  $|\Phi_g| \geq \|g\|_q$ , ce qui conclurait que  $|\Phi_g| = \|g\|_q$ . Nous posons  $f = g|g|^{q-2}$ , de telle sorte que  $|f| = |g|^{q-1}$  et

$$\|f\|_p = \left( \int |g|^{p(q-1)} \right)^{1/p} = \left( \int |g|^q \right)^{1/p} = \|g\|_q^{q/p} \quad (27.841)$$

où nous avons utilisé le fait que  $p(q-1) = q$ . La fonction  $f$  est donc bien dans  $L^p$ . D'autre part,

$$\Phi_g(f) = \int f \bar{g} = \int g |g|^{q-2} \bar{g} = \int |g|^q = \|g\|_q^q. \quad (27.842)$$

Donc

$$\frac{|\Phi_g(f)|}{\|f\|_p} = \|g\|_q^{q - \frac{q}{p}} = \|g\|_q \quad (27.843)$$

où nous avons encore utilisé le fait que  $q - \frac{q}{p} = \frac{q(p-1)}{p} = 1$ .

(3) **Surjectif**

Soit  $\ell \in (L^p)'$ ; c'est une application  $\ell: L^p \rightarrow \mathbb{C}$  dont nous pouvons prendre la restriction à  $L^2$  parce que la proposition 27.53 nous indique que  $L^2 \subset L^p$ . Nous nommons  $\phi: L^2 \rightarrow \mathbb{C}$  cette restriction.

(3a)  $\phi \in (L^2)'$

Nous devons montrer que  $\phi$  est continue pour la norme sur  $L^2$ . Pour cela nous montrons que sa norme opérateur (subordonnée à la norme de  $L^2$  et non de  $L^p$ ) est finie :

$$\sup_{f \in L^2} \frac{|\phi(f)|}{\|f\|_2} \leq \sup_{f \in L^2} \frac{|\ell(f)|}{\|f\|_p} < \infty. \quad (27.844)$$

Nous avons utilisé l'inégalité de norme  $\|f\|_p \leq \|f\|_2$  de la proposition 27.53(2).

(3b) **Utilisation du dual de  $L^2$**

Étant donné que  $L^2$  est un espace de Hilbert (lemme 27.105) et que  $\phi \in (L^2)'$ , le théorème 25.18 nous donne un élément  $g \in L^2$  tel que  $\phi(f) = \Phi_g(f)$  pour tout  $f \in L^2$ .

Nous devons prouver que  $g \in L^q$  et que pour tout  $f \in L^p$  nous avons  $\ell(f) = \Phi_g(f)$ .

(3c)  $g \in L^q$

Nous posons  $f_n = g|g|^{q-2} \mathbb{1}_{|g| < n}$ . Nous avons d'une part

$$\Phi_g(f_n) = \int_0^1 f_n \bar{g} = \int_{|g| < n} |g|^q. \quad (27.845)$$

Et d'autre part comme  $f_n \in L^2$  nous avons aussi  $\phi(f_n) = \Phi_g(f_n)$  et donc

$$0 \leq \Phi(f_n) = \phi(f_n) \leq \|\ell\| \|f_n\|_p \quad (27.846a)$$

$$= \|\ell\| \left( \int_{|g| < n} |g|^{(q-1)p} \right)^{1/p} \quad (27.846b)$$

$$= \|\ell\| \left( \int_{|g| < n} |g|^q \right)^{1/p}. \quad (27.846c)$$

où nous avons à nouveau tenu compte du fait que  $p(q-1) = q$ . En combinant avec (27.845) nous trouvons

$$\int_{|g|<n} |g|^q \leq \|\ell\| \left( \int_{|g|<n} |g|^q \right)^{1/p}, \quad (27.847)$$

et donc

$$\left( \int_{|g|<n} |g|^q \right)^{1-\frac{1}{p}} \leq \|\ell\|, \quad (27.848)$$

c'est-à-dire

$$\left( \int_{|g|<n} |g|^q \right)^{1/q} \leq \|\ell\|. \quad (27.849)$$

Si ce n'était pas encore fait nous nous fixons un représentant de la classe  $g$  (qui est dans  $L^2$ ), et nous nommons également  $g$  ce représentant. Nous posons alors

$$g_n = |g|^q \mathbb{1}_{|g|<n} \quad (27.850)$$

qui est une suite croissante de fonctions convergeant ponctuellement vers  $|g|^q$ . Le théorème de Beppo-Levi 14.178 nous permet alors d'écrire

$$\lim_{n \rightarrow \infty} \int_{|g|<n} |g|^q = \int_0^1 |g|^q. \quad (27.851)$$

Mais comme pour chaque  $n$  nous avons  $\int_{|g|<n} |g|^q \leq \|\ell\|^q$ , nous conservons l'inégalité à la limite et

$$\int_0^1 |g|^q \leq \|\ell\|^q. \quad (27.852)$$

Cela prouve que  $g \in L^p$ .

(3d)  $\ell(f) = \Phi_g(f)$

Soit  $f \in L^p$ . En vertu de la densité de  $L^2$  dans  $L^p$  prouvée dans le corolaire 27.66 nous pouvons considérer une suite  $(f_n)$  dans  $L^2$  telle que  $f_n \xrightarrow{L^p} f$ . Pour tout  $n$  nous avons

$$\ell(f_n) = \Phi_g(f_n). \quad (27.853)$$

Mais  $\ell$  et  $\Phi_g$  étant continues sur  $L^p$  nous pouvons prendre la limite et obtenir

$$\ell(f) = \Phi_g(f). \quad (27.854)$$

□

LemHNPEooHtMOGY

**Lemme 27.186** ([13]).

Soit  $(\Omega, \mathcal{A}, \mu)$  un espace mesuré fini. Soit  $g \in L^1(\Omega)$  et  $S$  fermé dans  $\mathbb{C}$ . Si pour tout  $E \in \mathcal{A}$  nous avons

$$\frac{1}{\mu(E)} \int_E g d\mu \in S, \quad (27.855)$$

alors  $g(x) \in S$  pour presque tout  $x \in \Omega$ .

*Démonstration.* Soit  $D = \overline{B(a, r)}$  un disque fermé dans le complémentaire de  $S$  (ce dernier étant fermé, le complémentaire est ouvert). Posons  $E = g^{-1}(D)$ . Prouvons que  $\mu(E) = 0$  parce que cela prouverait que  $g(x) \in D$  pour seulement un ensemble de mesure nulle. Mais  $S^c$  pouvant être écrit comme une union dénombrable de disques fermés<sup>103</sup>, nous aurions  $g(x) \in S^c$  presque nulle part.

<sup>103</sup>. Tout ouvert peut être écrit comme union dénombrable d'éléments d'une base de topologie par la proposition 7.2 et  $\mathbb{C}$  a une base dénombrable de topologie par la proposition 7.188.

Vu que  $\frac{1}{\mu(E)} \int_E a = a$  nous avons

$$\left| \frac{1}{\mu(E)} \int_E g d\mu - a \right| = \left| \frac{1}{\mu(E)} \int_E (g - a) \right| \leq \frac{1}{\mu(E)} \int_E |g - a| \leq \frac{1}{\mu(E)} \mu(E) r = r. \quad (27.856a)$$

Donc

$$\frac{1}{\mu(E)} \int_E g d\mu \in D, \quad (27.857)$$

ce qui est une contradiction avec le fait que  $D \subset S^c$ .  $\square$

Dans toute la partie d'analyse fonctionnelle, sauf mention du contraire, nous considérons dans  $L^p$  des fonctions à valeurs complexes, et donc les éléments du dual sont des applications linéaires continues à valeurs dans  $\mathbb{C}$ . La raison est que nous allons utiliser les résultats concernant  $L^p$  dans la partie sur les transformations de Fourier, tandis que les transformations de Fourier demandent naturellement de travailler sur les complexes.

**Théorème 27.187** (Théorème de représentation de Riesz, thème 21, [13, 521, 766, 507]).

Soit un espace mesuré  $(\Omega, \mathcal{A}, \mu)$ . Soit  $q$  tel que  $\frac{1}{p} + \frac{1}{q} = 1$  avec la convention que  $q = \infty$  si  $p = 1$ . Alors l'application

$$\begin{aligned} \Phi: L^q &\rightarrow (L^p)' \\ \Phi_g(f) &= \int_{\Omega} f \bar{g} d\mu \end{aligned} \quad (27.858)$$

est une bijection isométrique<sup>104</sup> dans les cas suivants :

(1) si  $1 < p < \infty$  et  $(\Omega, \mathcal{A}, \mu)$  est un espace mesuré quelconque,

(2) si  $p = 1$  et  $(\Omega, \mathcal{A}, \mu)$  est  $\sigma$ -fini.

*Démonstration.* Par petits bouts.

(1)  **$\Phi$  est injective** Nous commençons par prouver que  $\Phi$  est injectif. Soient  $g, g' \in L^q$  tels que  $\Phi_g = \Phi_{g'}$ . Alors pour tout  $f \in L^p$  nous avons

$$\int_{\Omega} f(\bar{g} - \bar{g}') d\mu = 0. \quad (27.859)$$

Soient des parties  $A_i$  de mesures finies telles que  $\Omega = \bigcup_{i=1}^{\infty} A_i$ . Étant donné que  $\mu(A_i)$  est fini, nous avons  $\mathbb{1}_{A_i} \in L^p(\Omega)$  et donc

$$\int_{A_i} (\bar{g} - \bar{g}') d\mu = \int_{\Omega} \mathbb{1}_{A_i}(x) (\bar{g} - \bar{g}')(x) d\mu(x) = 0. \quad (27.860)$$

La proposition 27.26 nous dit alors que  $g - g' = 0$  dans  $L^q(A_i)$ . Pour chaque  $i$ , la partie  $N_i = \{x \in A_i \text{ tel que } (g - g')(x) \neq 0\}$  est de mesure nulle.

Vu que  $\Omega$  est l'union de tous les  $A_i$ , la partie de  $\Omega$  sur laquelle  $g - g'$  est non nulle est l'union des  $N_i$  et donc de mesure nulle parce que une réunion dénombrable de parties de mesure nulle est de mesure nulle. Donc  $g - g' = 0$  presque partout dans  $\Omega$ , ce qui signifie  $g - g' = 0$  dans  $L^q(\Omega)$ .

(2) **La suite**

La partie difficile est de montrer que  $\Phi$  est surjective.

Soit  $\phi \in (L^p(\Omega))'$ . Si  $\phi = 0$ , c'est bien dans l'image de  $\Phi$ ; nous supposons donc que non. Nous allons commencer par prouver qu'il existe une (classe de) fonction  $g \in L^1(\Omega)$  telle que  $\Phi_g(f) = \phi(f)$  pour tout  $f \in L^{\infty}(\Omega, \mu)$ ; nous montrerons ensuite que  $g \in L^q$  et que le tout est une isométrie.

104. Pour rappel, la norme sur le dual est la norme opérateur 10.63.

(3) **Une mesure complexe**

Si  $E \in \mathcal{A}$  nous notons  $\nu(E) = \phi(\mathbb{1}_E)$ . Nous prouvons maintenant que  $\nu$  est une mesure complexe<sup>105</sup> sur  $(\Omega, \mathcal{A})$ . Le seul point pas facile est de démontrer est l'additivité dénombrable. Il est déjà facile de voir que  $A$  et  $B$  sont disjoints,  $\nu(A \cup B) = \nu(A) + \nu(B)$ . Soient ensuite des ensembles  $A_n$  deux à deux disjoints et posons  $E_k = \bigcup_{i \leq k} A_i$  pour avoir  $\bigcup_k A_k = \bigcup_k E_k$  avec l'avantage que les  $E_k$  soient emboîtés. Cela donne

$$\|\mathbb{1}_E - \mathbb{1}_{E_k}\|_p = \mu(E \setminus E_k)^{1/p}, \quad (27.861)$$

mais vu que  $1 \leq p < \infty$ , avoir  $x_k \rightarrow 0$  implique d'avoir  $x_k^{1/p} \rightarrow 0$ . Prouvons que  $\mu(E \setminus E_k) \rightarrow 0$ . En vertu du lemme 14.23 nous avons pour chaque  $k$  :

$$\mu(E \setminus E_k) = \mu(E) - \mu(E_k), \quad (27.862)$$

et vu que  $E_k \rightarrow E$  est une suite croissante, le lemme 14.24(1), sachant que  $\mu$  est une mesure « normale », donne

$$\lim_{n \rightarrow \infty} \mu(E_k) = \mu\left(\bigcup_k E_k\right). \quad (27.863)$$

Donc effectivement  $\mu(E_k) \rightarrow \mu(E)$  et donc oui :  $\mu(E \setminus E_k) \rightarrow 0$ . Jusqu'à présent nous avons

$$\lim_{k \rightarrow \infty} \|\mathbb{1}_E - \mathbb{1}_{E_k}\|_p = 0, \quad (27.864)$$

c'est-à-dire  $\mathbb{1}_{E_k} \xrightarrow{L^p} \mathbb{1}_E$ . La continuité de  $\phi$  sur  $L^p$  donne alors

$$\lim_{k \rightarrow \infty} \nu(E_k) = \lim_{k \rightarrow \infty} \phi(\mathbb{1}_{E_k}) = \phi\left(\lim_{k \rightarrow \infty} \mathbb{1}_{E_k}\right) = \phi(\mathbb{1}_E) = \nu(E). \quad (27.865)$$

Par additivité finie de  $\nu$  nous avons

$$\nu(E_k) = \sum_{i \leq k} \nu(A_i) \quad (27.866)$$

et en passant à la limite,  $\sum_{i=1}^{\infty} \nu(A_i) = \nu(\bigcup_i A_i)$ . L'application  $\nu$  est donc une mesure complexe.

(4) **Mesure absolument continue**

En prime, si  $\mu(E) = 0$  alors  $\nu(E) = 0$  parce que

$$\mu(E) = 0 \Rightarrow \|\mathbb{1}_E\|_p = 0 \Rightarrow \mathbb{1}_E = 0 \text{ (dans } L^p) \Rightarrow \phi(\mathbb{1}_E) = 0 \quad (27.867)$$

(5) **Utilisation de Radon-Nikodym**

Nous sommes donc dans un cas où  $\nu \ll \mu$  et nous utilisons le théorème de Radon-Nikodym 14.240 sous la forme de la remarque 14.241 : il existe une fonction intégrable  $g: \Omega \rightarrow \mathbb{C}$ <sup>106</sup> telle que pour tout  $A \in \mathcal{A}$ ,

$$\nu(A) = \int_A \bar{g} d\mu. \quad (27.868)$$

C'est-à-dire que

$$\phi(\mathbb{1}_A) = \int_A \bar{g} d\mu = \int_{\Omega} \bar{g} \mathbb{1}_A d\mu. \quad (27.869)$$

Nous avons donc exprimé  $\phi$  comme une intégrale pour les fonctions caractéristiques d'ensembles.

(6) **Pour les fonctions étagées**

Par linéarité si  $f$  est mesurable et étagée nous avons aussi

$$\phi(f) = \int f \bar{g} d\mu = \Phi_g(f). \quad (27.870)$$

105. Définition 14.239.

106. On peut écrire, pour utiliser de la notation compacte que  $g \in L^1(\Omega, \mathbb{C})$ .

(7) **Pour**  $f \in L^\infty(\Omega)$ 

Une fonction  $f \in L^\infty$  est une fonction presque partout bornée. Nous supposons que  $f$  est presque partout bornée par  $M$ . Par ailleurs cette  $f$  est limite uniforme de fonctions étagées :  $\|f_k - f\|_\infty \rightarrow 0$  en posant  $f_k = f \mathbb{1}_{|f| \leq k}$ . Pour chaque  $k$  nous avons l'égalité

$$\Phi_g(f_k) = \phi(f_k). \quad \text{EqPDCJooGNjwAQ} \quad (27.871)$$

Par ailleurs la fonction  $f_k \bar{g}$  est majorée par la fonction intégrable  $M \bar{g}$  et le théorème de la convergence dominée 14.207 nous donne

$$\lim_{k \rightarrow \infty} \Phi_g(f_k) = \lim_{k \rightarrow \infty} \int f_k \bar{g} = \int f \bar{g} = \Phi_g(f). \quad (27.872)$$

Et la continuité de  $\phi$  sur  $L^p$  couplée à la convergence  $f_k \xrightarrow{L^p} f$  donne  $\lim_{k \rightarrow \infty} \phi(f_k) = \phi(f)$ . Bref prendre la limite dans (27.871) donne

$$\Phi_g(f) = \phi(f) \quad (27.873)$$

pour tout  $f \in L^\infty(\Omega)$ .

(8) **La suite ...**

Voici les prochaines étapes.

- Nous avons  $\int f \bar{g} = \phi(f)$  tant que  $f \in L^\infty$ . Nous allons étendre cette formule à  $f \in L^p$  par densité. Cela terminera de prouver que notre application est une bijection.
- Ensuite nous allons prouver que  $\|\phi\| = \|\Phi_g\|$ , c'est-à-dire que la bijection est une isométrie.

(9) **De**  $L^\infty$  **à**  $L^p$ 

Soit  $f \in L^p$ . Si nous avons une suite  $(f_n)$  dans  $L^\infty$  telle que  $f_n \xrightarrow{L^p} f$  alors  $\lim \phi(f_n) = \phi(f)$  par continuité de  $\phi$ . La difficulté est de trouver une telle suite de façon à pouvoir permuter l'intégrale et la limite :

$$\lim_{n \rightarrow \infty} \int_\Omega f_n \bar{g} = \int_\Omega \lim_{n \rightarrow \infty} f_n \bar{g} = \int_\Omega f \bar{g} = \Phi_g(f). \quad \text{EqLYYAooUQnbfV} \quad (27.874)$$

Nous allons donc maintenant nous atteler à la tâche de trouver  $f_n \in L^\infty$  avec  $f_n \xrightarrow{L^p} f$  et telle que (27.874) soit valide.

Nous allons d'abord supposer que  $f \in L^p$  est positive à valeurs réelles. Nous avons alors par le théorème 14.118 qu'il existe une suite croissante de fonction étagées (et donc  $L^\infty$ ) telles que  $f_n \rightarrow f$  ponctuellement. De plus étant donné que  $|f_n| \leq |f|$ , la proposition 27.25 nous dit que  $f_n \xrightarrow{L^p} f$ . Pour chaque  $n$  nous avons

$$\int_\Omega f_n \bar{g} = \phi(f_n). \quad (27.875)$$

Soit  $g^+$  la partie réelle positive de  $\bar{g}$ . Alors nous avons la limite croissante ponctuelle  $f_n g^+ \rightarrow f g^+$  et le théorème de la convergence monotone 14.178 nous permet d'écrire

$$\lim_{n \rightarrow \infty} \int f_n g^+ = \int f g^+. \quad (27.876)$$

Faisant cela pour les trois autres parties de  $\bar{g}$  nous avons prouvé que si  $f \in L^p$  est réelle et positive,

$$\int f \bar{g} = \phi(f), \quad (27.877)$$

c'est-à-dire que  $\Phi_g(f) = \phi(f)$ .

Refaisant le tout pour les trois autres parties de  $f$  nous montrons que

$$\Phi_g(f) = \phi(f) \quad (27.878)$$

pour tout  $f \in L^p(\Omega)$ . Nous avons donc égalité de  $\phi$  et  $\Phi_g$  dans  $(L^p)'$  et donc bijection entre  $(L^p)'$  et  $L^q$ .



(10) **Isométrie : mise en place**

Nous devons prouver que cette bijection est isométrique. Soit  $\phi \in (L^p)'$  et  $g \in L^q$  telle que  $\Phi_g = \phi$ . Il faut prouver que

$$\|g\|_q = \|\phi\|_{(L^p)'} . \quad (27.879)$$

(11)  $\|\phi\| \leq \|g\|_q$ 

Nous savons que  $\phi(f) = \int f \bar{g}$ , et nous allons écrire la définition de la norme dans  $(L^p)'$  :

$$\|\phi\|_{(L^p)'} = \sup_{\|f\|_p=1} |\phi(f)| \quad (27.880a)$$

$$= \sup \left| \int f \bar{g} \right| \quad (27.880b)$$

$$\leq \sup \underbrace{\left| \int f \bar{g} \right|}_{=\|f \bar{g}\|_1} . \quad (27.880c)$$

Il s'agit maintenant d'utiliser l'inégalité de Hölder 27.39 :

$$\|\phi\| \leq \sup_{\|f\|_p=1} \|f\|_p \|\bar{g}\|_q = \|g\|_q . \quad (27.881)$$

L'inégalité dans l'autre sens sera démontrée en séparant les cas  $p = 1$  et  $1 < p < \infty$ .

(12) **Si  $p = 1$ , une formule** Si  $E$  est un ensemble mesurable de mesure finie, alors

$$\left| \int_E g d\mu \right| = |\phi(\mathbb{1}_E)| . \quad (27.882)$$

Mais le fait que  $\mu(E) < \infty$  donne que  $\mathbb{1}_E \in L^1(\Omega)$ . Donc  $\mathbb{1}_E \in L^\infty \cap L^1$ ; nous pouvons alors écrire  $\phi(\mathbb{1}_E) = \int_\Omega \mathbb{1}_E \bar{g} d\mu$  et donc

$$\left| \int_\Omega \mathbb{1}_E \bar{g} d\mu \right| = \left| \int_E \bar{g} d\mu \right| = |\phi(\mathbb{1}_E)| \leq \|\phi\|_{(L^1)'} \|\mathbb{1}_E\|_1 = \|\phi\| \mu(E) . \quad (27.883) \quad \text{EqUPCTooIvoKKI}$$

Nous écrivons cela dans l'autre sens :

$$\|\phi\| \geq \frac{1}{\mu(E)} \left| \int_\Omega \mathbb{1}_E \bar{g} d\mu \right| = \frac{1}{\mu(E)} \left| \int_E \bar{g} d\mu \right| . \quad (27.884)$$

Si nous prenons  $S = \{t \in \mathbb{C} \text{ tel que } |t| \leq \|\phi\|\}$ , c'est un fermé vérifiant que

$$\frac{1}{\mu(E)} \int_E \bar{g} d\mu \in S . \quad (27.885) \quad \text{EqQoMRLGooYPEjUo}$$

Voilà une petite formule qui va nous aider à utiliser le lemme 27.186. Nous ne pouvons cependant pas l'utiliser immédiatement parce que l'appartenance (27.885) n'est vraie que pour les parties de mesure finie.

(13) **Si  $p = 1$ , conclusion[1]**

Pour utiliser le lemme 27.186, nous utilisons l'hypothèse que  $\Omega$  est  $\sigma$ -fini. Soient des mesurables  $A_i$  de mesure finie tels que  $\bigcup_{i \in \mathbb{N}} A_i = \Omega$ .

Pour chaque  $i$  nous considérons la restriction  $g_i: A_i \rightarrow \mathbb{C}$  de  $g$  à  $A_i$ . Par le point précédent, elle vérifie

$$\frac{1}{\mu(A_i)} \int_{A_i} \bar{g}_i d\mu = \frac{1}{\mu(A_i)} \int_{A_i} \bar{g} d\mu \in S . \quad (27.886)$$

En appliquant le lemme 27.186 à l'espace restreint  $(A_i, \mathcal{A}_i, \mu_i)$ , nous concluons  $\bar{g}_i \in S$  presque partout, ce qui signifie que  $\|g_i\|_\infty \in S$ . Nous en concluons que

$$\|g_i\|_\infty \leq \|\phi\| \quad (27.887)$$

où, dans ce contexte,  $\|g_i\|_\infty$  signifie  $\sup_{x \in A_i} |g_i(x)|$ .

Nous avons alors

$$\|g\|_\infty = \sup_{x \in \Omega} |g(x)| = \sup_{i \in \mathbb{N}} \|g_i\|_\infty \leq \|\phi\|. \quad (27.888)$$

Une petite justification pour cela ? Prenons une suite  $x_k$  telle que  $|g(x_k)| \rightarrow \|g\|_\infty$ . Vu que les  $A_i$  recouvrent  $\Omega$ , existe un naturel  $i(k)$  tel que  $x_k \in A_{i(k)}$ . Nous avons alors

$$|g(x_k)| \leq \|g_{i(k)}\|_\infty \leq \|\phi\|. \quad (27.889)$$

Cela pour conclure que  $g \in L^\infty$ .

Notons que cet argument ne tient pas avec  $p > 1$  parce que l'équation (27.883) terminerait sur  $\|\phi\|\mu(E)^{1/p}$ . Du coup l'ensemble  $S$  à prendre serait  $S = \{t \in \mathbb{C} \text{ tel que } |t| \leq \|\phi\|\mu(E)^{1/p-1}\}$  et nous sommes en dehors des hypothèses du lemme parce qu'il n'y a pas d'ensemble *indépendant* de  $E$  dans lequel l'intégrale  $\frac{1}{\mu(E)} \int_E \bar{g} d\mu$  prend ses valeurs.

(14)  $1 < p < \infty$

La fonction

$$\alpha(x) = \begin{cases} \frac{g(x)}{|g(x)|} & \text{si } g(x) \neq 0 \\ 1 & \text{si } g(x) = 0 \end{cases} \quad (27.890)$$

a la propriété de faire  $\alpha g = |g|$  en même temps que  $|\alpha(x)| = 1$  pour tout  $x$ . Nous définissons

$$E_n = \{x \text{ tel que } |g(x)| \leq n\} \quad (27.891)$$

et

$$f_n = \mathbb{1}_{E_n} |g|^{q-1} \alpha. \quad (27.892)$$

Ce qui est bien avec ces fonctions c'est que<sup>107</sup>

$$|f_n|^p = |g|^{p(q-1)} |\alpha|^p = |g|^q \quad (27.893)$$

sur  $E_n$ . Dans  $E_n$  nous avons  $|f_n| = |g|^{q-1} \leq n^{q-1}$  et dans  $E_n$  nous avons  $f_n = 0$ . Au final,  $f_n \in L^\infty$ . Par ce que nous avons vu plus haut, nous avons alors

$$\phi(f_n) = \Phi_g(f_n). \quad (27.894)$$

Par ailleurs,

$$f_n \bar{g} = \mathbb{1}_{E_n} |g|^{q-1} \frac{g}{|g|} \bar{g}, \quad (27.895)$$

donc<sup>108</sup>

$$\left| \int_{E_n} |g|^q d\mu \right| = \left| \int_{E_n} f_n \bar{g} d\mu \right| \quad (27.896a)$$

$$= |\phi(f_n)| \quad (27.896b)$$

$$\leq \|\phi\| \|f_n\|_p \quad (27.896c)$$

$$= \|\phi\| \left( \int_{E_n} |f_n|^p \right)^{1/p} \quad (27.896d)$$

$$= \|\phi\| \left( \int_{E_n} |g|^q \right)^{1/p}. \quad (27.896e)$$

Nous avons de ce fait une inégalité de la forme  $A \leq \|\phi\| A^{1/p}$  et donc aussi  $A^{1/p} \leq \|\phi\|^{1/p} A^{1/p^2}$ , et donc  $A \leq \|\phi\| \|\phi\|^{1/p} A^{1/p^2}$ . Continuant ainsi à injecter l'inégalité dans elle-même, pour tout  $k \in \mathbb{N}$  nous avons :

$$\left| \int_{E_n} |g|^q d\mu \right| \leq \|\phi\|^{1+\frac{1}{p}+\dots+\frac{1}{p^k}} \left( \int_{E_n} |g|^q d\mu \right)^{1/p^k}. \quad (27.897)$$

107. C'est ici que nous utilisons le lien entre  $p$  et  $q$ . En l'occurrence, de  $1/p + 1/q = 1$  nous déduisons  $q(p-1) = p$ .

108. Dans [13], cette équation arrive sans modules, ce qui me laisse entendre que  $\phi(f_n)$  est réel et positif pour pouvoir écrire que  $\phi(f_n) \leq \|\phi\| \|f_n\|_p$ , mais je ne comprends pas pourquoi.

Nous pouvons passer à la limite  $k \rightarrow \infty$ . Sachant que  $p > 1$  nous savons  $A^{1/k} \rightarrow 1$  et

$$1 + \frac{1}{p} + \cdots + \frac{1}{p^k} \rightarrow \frac{p}{p-1} = q. \quad (27.898)$$

Nous avons alors

$$\int_{E_n} |g|^q d\mu \leq \|\phi\|^q. \quad (27.899)$$

L'intégrale s'écrit tout aussi bien sous la forme  $\int_{\Omega} |g|^q \mathbb{1}_{E_n}$ . La fonction dans l'intégrale est une suite croissante de fonctions mesurables à valeurs dans  $[0, \infty]$ . Nous pouvons alors permuter l'intégrale et la limite  $n \rightarrow \infty$  en utilisant la convergence monotone (théorème 14.178) qui donne alors  $\int_{\Omega} |g|^q \leq \|\phi\|^q$  ou encore

$$\|g\|_q \leq \|\phi\|. \quad (27.900)$$

Ceci achève de prouver que l'application  $\phi \mapsto \Phi_g$  est une isométrie, et donc le théorème. □

THOooXmVTooBAbyvr

### **Théorème 27.188.**

Soit un espace mesuré  $(\Omega, \mathcal{A}, \mu)$ . Pour  $f, g \in \text{Fun}(\Omega, \mathbb{C})$ , nous écrivons

$$\Phi_g(f) = \int_{\Omega} f \bar{g} \quad (27.901)$$

pour toutes les combinaisons de  $f$  et  $g$  pour lesquelles l'intégrale a un sens.

ITEMooNCVEooTyNsoJ

(1) Si  $1 < p < \infty$ , alors  $L^p(\Omega, \mathcal{A}, \mu)$  est réflexif<sup>109</sup>.

ITEMooTQDJooFShTiA

(2) Si  $(\Omega, \mathcal{A}, \mu)$  est  $\sigma$ -finie, alors

ITEMooHMMZooMQxWgB

(2a) L'application  $\Phi: L^\infty(\Omega, \mathcal{A}, \mu) \rightarrow L^1(\Omega, \mathcal{A}, \mu)'$  est une bijection isométrique.

ITEMooBFFZooNxoHER

(2b) L'application  $\Phi: L^1(\Omega, \mathcal{A}, \mu) \rightarrow L^\infty(\Omega, \mathcal{A}, \mu)'$  est une injection isométrique.

Note : nous verrons dans 27.192 que l'application  $\Phi: L^1(\mathbb{R}^d) \rightarrow L^\infty(\mathbb{R}^d)'$  n'est pas surjective.

*Démonstration.* En plusieurs parties, en notant toujours  $p$  et  $q$  les exposants conjugués, c'est-à-dire  $\frac{1}{p} + \frac{1}{q} = 1$ .

(1) **Pour (1)**

Nous considérons les applications

$$\begin{aligned} \alpha_1: L^p &\rightarrow (L^q)' \\ \alpha_1(g)f &= \int_{\Omega} f \bar{g} \end{aligned} \quad (27.902)$$

et

$$\begin{aligned} \alpha_2: L^q &\rightarrow (L^p)' \\ \alpha_2(g)f &= \int_{\Omega} g \bar{f}. \end{aligned} \quad (27.903)$$

Ce sont des bijections linéaires par le théorème de représentation de Riesz 27.187<sup>110</sup>. L'espace  $L^p$  est de Banach par le théorème de Riesz-Fischer 27.56; le lemme 27.183 s'applique donc. Nous considérons alors les applications correspondantes

$$\beta_1: (L^p)' \rightarrow (L^q)'' \quad (27.904)$$

et

$$\beta_2: (L^q)' \rightarrow (L^p)'', \quad (27.905)$$

109. Définition 27.179.

110. Notez que la conjugaison complexe dans  $\alpha_2$  n'est pas à la même place que dans  $\alpha_1$ . L'application  $\alpha_1$  est exactement la même que le  $\Phi$  de représentation de Riesz alors que  $\alpha_2$  est un tout petit peu modifiée. La raison de ce changement n'est pas très profonde : c'est seulement pour que les choses tombent juste à la fin.

qui sont également des bijections linéaires. Nous allons montrer que la bijection

$$\beta_2 \circ \alpha_1: L^p \rightarrow (L^p)'' \quad (27.906)$$

est l'application  $J$  de la définition 27.179.

Soient  $f \in L^p$  et  $\varphi \in (L^p)'$ . Il existe  $g \in L^q$  tel que  $\varphi = \alpha_2(g)$ . Nous pouvons donc calculer d'une part, en développant  $\beta_2$  par sa définition (27.827),

$$(\beta_2 \circ \alpha_1)(f)\varphi = \beta_2(\alpha_1(f))\varphi = \alpha_1(f)(\alpha_2^{-1}(\varphi)) = \alpha_1(f)(g) = \int_{\Omega} g\bar{f}, \quad (27.907)$$

et d'autre part

$$J(f)\varphi = \varphi(f) = \alpha_2(g)f = \int_{\Omega} g\bar{f}. \quad (27.908)$$

(2) **Pour (2a)** Il s'agit du théorème 27.187(2).

(3) **Pour (2b)[1]** Il nous reste à couvrir le cas de  $(L^\infty)'$ . Pour  $g \in L^1$  nous prouvons que  $\Phi_g \in (L^\infty)'$ .

(3a)  **$\Phi_g(f)$  est bien définie** Nous prouvons d'abord que si  $f \in L^\infty$ , alors l'intégrale  $\int_{\Omega} f\bar{g}$  est bien définie. Par définition du supremum essentiel <sup>111</sup>, il existe  $M > 0$  tel que  $|f(x)| < M$  pour tout  $x$  hors d'une partie  $A$  de mesure nulle. Nous avons alors

$$\int_{\Omega} |f\bar{g}| = \int_{\Omega \setminus A} |f\bar{g}| \leq M \int_{\Omega \setminus A} |g| = M \int_{\Omega} |g| < \infty. \quad (27.909)$$

(3b)  **$\Phi_g$  est continue** Soit une suite  $f_k \xrightarrow{L^\infty} f$  ainsi que  $g \in L^1$ . Pour chaque  $k$ , il existe une partie de mesure nulle  $A_k$  et un nombre  $M_k = \|f_k\|_{L^\infty}$  tel que  $|f_k(x)| < M_k$  pour tout  $x$  hors de  $A_k$ . Nous avons alors

$$|\Phi_g(f_k)| \leq \int_{\Omega \setminus A_k} |f_k\bar{g}| d\mu \leq \|f_k\|_{L^\infty} \int_{\Omega \setminus A_k} |g| \leq \|f_k\|_{L^\infty} \|g\|_1. \quad (27.910)$$

Vu que par hypothèse  $f_k \rightarrow 0$  dans  $L^\infty$ , nous avons  $\|f_k\|_{L^\infty} \rightarrow 0$ , et donc aussi

$$|\Phi_g(f_k)| \rightarrow 0. \quad (27.911)$$

□

## 27.189.

Les gens qui n'ont peur d'aucun abus de notations écrivent le théorème 27.188(2a) en disant simplement que  $L^1 = L^\infty$  et le démontrent de la façon suivante. Le théorème 27.187(1) nous indique que

$$(L^p)' = L^q. \quad (27.912)$$

Vu que  $1 < p < \infty$ , nous avons aussi  $1 < q < \infty$  et donc  $(L^q)' = L^p$ . En prenant le dual des deux côtés de (27.912),

$$(L^p)'' = (L^q)' = L^p. \quad (27.913)$$

À ce moment, un second abus vient en aide et nous disons qu'un espace  $V$  est réflexif quand  $V = V''$ . Et voilà.

## Proposition 27.190.

Soit  $f \in L^p(\Omega)$  telle que

$$\int_{\Omega} f\varphi = 0 \quad (27.914)$$

pour tout  $\varphi \in \mathcal{D}(\Omega)$ . Alors  $f = 0$  presque partout.

<sup>111</sup>. Voir les définitions 27.28 et 27.29.

*Démonstration.* Nous considérons la forme linéaire  $\Phi_f \in (L^q)'$  donnée par

$$\begin{aligned}\Phi_f: L^p &\rightarrow \mathbb{C} \\ u &\mapsto \int_{\Omega} f u\end{aligned}\tag{27.915}$$

Par hypothèse cette forme est nulle sur la partie dense  $C_c^\infty(\Omega)$ . Si  $(\varphi_n)$  est une suite dans  $C_c^\infty(\Omega)$  convergente vers  $u$  dans  $L^p$ , nous avons pour tout  $n$  que

$$0 = \Phi_f(\varphi_n)\tag{27.916}$$

En passant à la limite, nous voyons que  $\Phi_f$  est la forme nulle. Elle est donc égale à  $\Phi_0$ . La partie « unicité » du théorème de représentation de Riesz 27.187 nous indique alors que  $f = 0$  dans  $L^p$  et donc  $f = 0$  presque partout.  $\square$

PropLGoLtcS

**Proposition 27.191.**

Si  $f \in L_{loc}^1(I)$  est telle que

$$\int_I f \varphi' = 0\tag{27.917}$$

pour tout  $\varphi \in C_c^\infty(I)$ , alors il existe une constante  $C$  telle que  $f = C$  presque partout.

*Démonstration.* Soit  $\psi \in C_c^\infty(I)$  une fonction d'intégrale 1 sur  $I$ . Si  $w \in C_c^\infty(I)$  alors nous considérons la fonction

$$h = w - \psi \int_I w,\tag{27.918}$$

qui est dans  $C_c^\infty(I)$  et dont l'intégrale sur  $I$  est nulle. Par la proposition 17.2, la fonction  $h$  admet une primitive dans  $C_c^\infty(I)$ ; et nous notons  $\varphi$  cette primitive. L'hypothèse appliquée à  $\varphi$  donne

$$0 = \int_I f \varphi' = \int_I f \left( w - \psi \int_I w \right) = \int_I f w - \underbrace{\left( \int_I f(x) \psi(x) dx \right)}_C \left( \int_I w(y) dy \right) = \int_I w(f - C).\tag{27.919}$$

L'annulation de la dernière intégrale implique par la proposition 27.190 que  $f - C = 0$  dans  $L^2$ , c'est-à-dire  $f = C$  presque partout.  $\square$

PROPooXNR RooUdgFP

**Proposition 27.192 ([220]).**

L'application

$$\begin{aligned}\Phi: L^1(\mathbb{R}^d) &\rightarrow L^\infty(\mathbb{R}^d)' \\ \Phi_g(f) &= \int_{\Omega} f \bar{g} d\lambda\end{aligned}\tag{27.920}$$

n'est pas surjective.

*Démonstration.* Nous allons construire un élément de  $(L^\infty)'$  qui n'est pas dans l'image de  $L^1$ .

- (1) **Mise en place du décor** Nous considérons l'espace vectoriel normé  $(L^\infty(\mathbb{R}^d), N_\infty)$  défini en 27.28 et 27.29. Dedans, nous considérons le sous-espace  $D$  des classes des fonctions dans  $\mathcal{D}(\mathbb{R}^d)$ ; nous aurions très envie de le noter  $(\mathcal{D}(\mathbb{R}^d), N_\infty)$ , mais nous allons le noter  $(D, N_\infty)$  parce que  $\mathcal{D}$  est un espace de fonctions tandis que nous considérons un espace de classes de fonctions.

Si  $f \in D$ , alors  $\tilde{f}$  est une classe de fonctions contenant un unique représentant continu<sup>112</sup>; nous le notons  $\tilde{f}$ .

<sup>112.</sup> Existence parce que les éléments de  $D$  sont des classes d'éléments de  $\mathcal{D}$  qui sont  $C^\infty$ . Unicité par la proposition 20.154.

(2) Une première fonctionnelle

Avec ça nous posons <sup>113</sup>

$$\begin{aligned}\phi_0: D &\rightarrow \mathbb{C} \\ f &\mapsto \tilde{f}(0).\end{aligned}\tag{27.921}$$

(3)  $\phi_0$  est linéaire Ça ne devrait pas poser de problèmes.

(4)  $\phi_0$  est continue Soit  $f_k \xrightarrow{N_\infty} 0$ . Nous devons prouver que  $\tilde{f}_k(0) \xrightarrow{\mathbb{C}} 0$ . Supposons le contraire et considérons  $\epsilon > 0$  ainsi que  $k$  tels que  $\tilde{f}_k(0) > \epsilon$ .

Par continuité de  $\tilde{f}_k$ , il existe un  $\delta > 0$  tel que  $\tilde{f}_k(x) > \epsilon/2$  pour tout  $x \in B(0, \delta)$ . Avec cela nous avons  $N_\infty(f_k) > \epsilon/2$ , et une impossibilité d'avoir <sup>114</sup>  $f_k \xrightarrow{N_\infty} 0$ .

(5)  $\phi_0$  est de norme finie C'est parce qu'elle est continue.

(6) Utilisation de Hahn-Banach Le théorème de Hahn-Banach 27.171 donne une extension continue

$$\phi: L^\infty(\mathbb{R}^d) \rightarrow \mathbb{C}.\tag{27.922}$$

Cela est un élément de  $(L^\infty)'$  et nous allons montrer qu'il n'est pas dans  $\Phi(L^1)$ .

(7) Par l'absurde Supposons qu'il existe  $u \in L^1(\mathbb{R}^d)$  telle que  $\phi = \Phi_u$ . Pour tout  $f \in L^\infty$ , nous avons

$$\phi(f) = \Phi_u(f) = \int_{\mathbb{R}^d} f \bar{u}.\tag{27.923}$$

Si  $f$  est la classe d'une fonction de  $\mathcal{D}(\mathbb{R}^d)$  s'annulant en 0, alors  $\phi(f) = 0$ . De telles fonctions non identiquement nulles existent par le lemme 15.176 <sup>115</sup>.

Soient  $\varphi \in \mathcal{D}(\mathbb{R}^d)$ , un compact  $K$  ne contenant pas 0 et  $\chi_K$  sa fonction indicatrice. Nous avons  $\chi_K \varphi \in \mathcal{L}^\infty(\mathbb{R}^d)$ , et  $(\chi_K \varphi)(0) = 0$ . En passant aux classes,  $\phi([\chi_K f]) = 0$ . Nous avons :

$$0 = \int_{\mathbb{R}^d} u \overline{(\chi_K \varphi)} = \int_{\mathbb{R}^d} \bar{\varphi} (u \overline{\chi_K}).\tag{27.924}$$

Vu que cela est valable pour tout  $\varphi \in \mathcal{D}(\mathbb{R}^d)$ , la proposition 27.190 dit que  $u \chi_K = 0$  presque partout.

(8) Le coup du compact

Soit une suite de compacts  $K_n$  recouvrant  $\mathbb{R}^d \setminus \{0\}$ . Par exemple

$$K_n = \overline{B(0, n)} \setminus B(0, \frac{1}{n}).\tag{27.925}$$

Pour chacun des  $K_n$ , nous avons  $u \chi_{K_n} = 0$  presque partout. Il existe donc une partie de mesure nulle  $N_k$  telle que  $u \chi_{K_n}$  est nulle à part sur  $N_k$ . Au total,  $u$  est non nulle seulement sur  $\bigcup_k N_k$  et peut-être en 0.

Bref,  $u$  est non nulle sur une partie de mesure nulle par le lemme 14.28.

(9) Conclusion

La fonction  $u$  est nulle presque partout. Donc  $\phi = \Phi_u = 0$ . Nous savons pourtant que  $\phi$  n'est pas nulle parce qu'il existe des fonctions dans  $\mathcal{D}(\mathbb{R}^d)$  qui ne s'annulent pas en 0.

Cela est une contradiction. Donc  $\phi$  n'est pas dans l'image de  $\Phi$  tout en étant dans  $L^\infty(\mathbb{R}^d)'$ .

□

Dans [767], il est dit que « la preuve [du lemme suivant], un peu fastidieuse mais en rien ingénieuse, est laissée en exercice ». La preuve est donc de moi ; elle est un tout petit peu ingénieuse mais en rien fastidieuse. J'espère ne pas m'être trompé et me demande bien ce que l'auteur avait en tête. Ma preuve s'appuie sur la proposition 17.101 dont la preuve ne me paraît pas non plus « fastidieuse mais en rien ingénieuse ».

113. Dans [220], l'auteur ne définit pas  $L^\infty$  comme un espace de classes de fonctions, et ces complications disparaissent.

114. Je vous laisse emballer ce raisonnement dans « si pour tout  $N$ , il existe  $k > N$  tel que ».

115. Ce lemme donne un exemple  $f$  sur  $\mathbb{R}$ . Si vous voulez vraiment un exemple dans  $\mathbb{R}^d$ , prenez  $g(x) = f(\|x\|)$ .

LEMooLDQRooEGWDlm

**Lemme 27.193** ([767, 1]).

Soit  $r > 0$ . Il existe  $\delta > 0$  tel que pour tout  $s, t \in \mathbb{C}$  vérifiant  $|s| \leq 1$ ,  $|t| \leq 1$  et  $|s - t| \geq r$  nous ayons

$$\left| \frac{s+t}{2} \right|^p \leq (1-\delta) \frac{|s|^p + |t|^p}{2}. \quad (27.926)$$

*Démonstration.* Soit  $r > 0$ . La partie de  $\mathbb{C}^2$  donnée par

$$D = \{(s, t) \in \mathbb{C}^2 \text{ tel que } |s| \leq 1, |t| \leq 1, |s - t| \geq r\} \quad (27.927)$$

est compacte. En effet elle est bornée (par la sphère de rayon  $\sqrt{2}$ ) et fermée comme intersection de fermés<sup>116</sup>. Nous considérons la fonction  $\Delta: D \rightarrow \mathbb{R}$  donnée par

$$\left| \frac{s+t}{2} \right|^p = \Delta(s, t) \frac{|s|^p + |t|^p}{2}. \quad (27.928)$$

Si vous voulez une expression explicite,

$$\Delta(s, t) = \frac{2^{p-1}|s+t|^p}{|s|^p + |t|^p}. \quad (27.929)$$

Cela est bien défini et continu sur  $D$  parce que le complémentaire  $D^c$  (qui est ouvert) contient  $(0, 0)$  et donc aussi un voisinage de  $(0, 0)$ .

La proposition 17.101 nous dit que la fonction  $z \mapsto |z|^p$  est strictement convexe. En prenant la définition 17.94 de la stricte convexité avec  $\theta = \frac{1}{2}$ , nous trouvons que

$$\Delta(s, t) < 1 \quad (27.930)$$

pour tout  $(s, t) \in D$ . Vu que par ailleurs  $\Delta$  est une fonction continue sur le compact  $D$ , elle atteint un minimum dans  $D$ . Soit  $\Delta_0$  ce minimum qui vérifie forcément  $\Delta_0 < 1$ .

En posant  $1 - \delta = \Delta_0$  nous avons le résultat.  $\square$

### 27.16.1 Le théorème de Jordan

Les définitions de chemins, de lacets et de courbes de Jordan sont dans 20.75.

LEMooZPRLooJPvrOE

**Lemme-Définition 27.194** ([234]).

Soit un lacet  $\gamma: [a, b] \rightarrow X$ . Nous considérons l'application

$$\begin{aligned} q: [a, b] &\rightarrow S^1 \\ t &\mapsto e^{2i\pi t/(b-a)} \end{aligned} \quad (27.931)$$

Il existe une unique application continue  $\tilde{\gamma}: S^1 \rightarrow \text{Image}(\gamma)$  telle que  $\gamma = \tilde{\gamma} \circ q$ .

Cette application  $\tilde{\gamma}$  est l'application **quotient** associée à  $\gamma$ .

*Démonstration.* La proposition 18.62(1) dit que  $q$  est une bijection continue. Donc l'existence et l'unicité d'une application  $\tilde{\gamma} = \gamma \circ q^{-1}$ . Cette application est continue comme composée d'applications continues.  $\square$

LEMooCGVOooVPISRD

**Lemme 27.195** ([234]).

Soit  $\gamma: [a, b] \rightarrow X$  un chemin d'image  $\Gamma$ .

ITEMooWKVAooCQDvpL

- (1) Si  $\gamma$  est un chemin de Jordan<sup>117</sup>, alors  $\gamma: [a, b] \rightarrow \Gamma$  est un isomorphisme d'espaces topologiques.

116. Lemme 7.6 suivit du théorème de Borel-Lebesgue 9.24.

117. Définition 20.75.

(2) Si  $\gamma$  est un lacet de Jordan, alors le quotient<sup>118</sup>  $\tilde{\gamma}: S^1 \rightarrow \Gamma$  est un homéomorphisme<sup>119</sup>. ITEMooVYMxooEtgPJT

*Démonstration.* En deux parties.

- (1) **Pour (1)** Un chemin de Jordan est injectif, et donc bijectif sur son image. Il est donc une bijection continue depuis  $[a, b]$  qui est compact. Il est donc un homéomorphisme par le lemme 7.276.
- (2) **Pour (2)** L'application  $\gamma: [a, b[ \rightarrow \Gamma$  est une bijection continue parce que  $\gamma$  est un lacet de Jordan. Son quotient  $\tilde{\gamma}: S^1 \rightarrow \Gamma$  est donc une bijection continue depuis le compact  $S^1$ . Le lemme 7.276 conclut que c'est un homéomorphisme.

□

CORooPGFLooUTVZMi

### Corolaire 27.196.

À propos de courbes simples et de Jordan.

- (1) Toute courbe simple est homéomorphe à  $[0, 1]$ .
- (2) Toute courbe de Jordan est homéomorphe à  $S^1$ .

ITEMooIUAXooOfvNov

*Démonstration.* Une courbe simple est l'image d'un lacet de Jordan. Donc  $\gamma: [a, b] \rightarrow \Gamma$  est un homéomorphisme par le lemme 27.195(1). Le lemme 10.295 fournit un homéomorphisme entre  $[a, b]$  et  $[0, 1]$ . Une composition d'homéomorphismes est un homéomorphisme.

Une courbe de Jordan est l'image d'un lacet de Jordan  $\gamma: [a, b] \rightarrow \Gamma$ . Le lemme 27.195(2) dit alors que  $\Gamma$  est homéomorphe à  $S^1$ . □

DEFooURFMooXIaRkl

### Définition 27.197 ([234]).

Soit un compact  $K$  dans  $\mathbb{C}$ . Nous notons

- (1)  $\mathcal{G}(K)$  le groupe multiplicatif des fonctions continues  $f: K \rightarrow \mathbb{C}^*$ .
- (2)  $\mathcal{E}(K)$  le sous-groupe des fonctions admettant un logarithme continu<sup>120</sup>.
- (3)  $G(K)$  le quotient<sup>121</sup>  $G(K) = \mathcal{G}(K)/\mathcal{E}(K)$ .

LEMooHEOWooHTtHsJ

### Lemme 27.198 ([234]).

Si  $K_1$  et  $K_2$  sont des compacts homéomorphes, alors les groupes  $G(K_1)$  et  $G(K_2)$  sont isomorphes.

*Démonstration.* Soit un homéomorphisme  $\theta: K_1 \rightarrow K_2$ . Nous posons

$$\begin{aligned} \psi: \mathcal{G}(K_1) &\rightarrow \mathcal{G}(K_2) \\ f &\mapsto f \circ \theta^{-1}. \end{aligned} \quad (27.932)$$

- (1) **Injective** Suppose  $\psi(f) = \psi(g)$ . Vu que  $\theta$  est une bijection, l'égalité  $f \circ \theta^{-1} = g \circ \theta^{-1}$  implique  $f = g$ . Donc  $\psi$  est injective.
- (2) **Surjective** Soit  $f \in \mathcal{G}(K_2)$ . Nous avons  $f \circ \theta \in \mathcal{G}(K_1)$ , et évidemment  $\psi(f \circ \theta) = f$ .
- (3) **Morphisme** Nous avons

$$\psi(fg) = fg \circ \theta^{-1} = (f \circ \theta^{-1})(g \circ \theta^{-1}) = \psi(f)\psi(g). \quad (27.933)$$

Nous avons prouvé que  $\psi$  est un isomorphisme. Nous devons maintenant voir qu'il passe au quotient, c'est-à-dire que  $\psi(\mathcal{E}(K_1)) \subset \mathcal{E}(K_2)$ . Soit  $f \in \mathcal{E}(K_1)$ , soit un logarithme continu  $g$  de  $f$ , c'est-à-dire  $\exp(g(x)) = f(x)$ . Pour  $y \in K_2$ , il existe  $x \in K_1$  tel que  $y = \theta(x)$ . Dans ce cas nous avons

$$\exp((g \circ \theta^{-1})(y)) = \exp(g(x)) = f(x) = (f \circ \theta^{-1})(y). \quad (27.934)$$

Autrement dit,

$$\exp(\psi(g)(y)) = \psi(f). \quad (27.935)$$

Donc  $\psi(g)$  est un logarithme continu de  $\psi(f)$ . □

118. Définition 27.194.

119. Bijection continue d'inverse continu.

120. Logarithme continu, définition 26.45.

121. Le groupe  $\mathcal{G}(K)$  est commutatif; donc pas de problèmes pour que  $\mathcal{E}(K)$  soit normal.



DEFooKXHUooTsKuOe

**Définition 27.199.**

*Petite notation. Si un compact  $K$  est donné, pour  $p \in \mathbb{C} \setminus K$ , nous notons*

$$\begin{aligned} f_p: K &\rightarrow \mathbb{C}^* \\ z &\mapsto z - p. \end{aligned} \quad (27.936)$$

**Lemme 27.200.**

LEMooSULYooWiEoyf

*Nous considérons l'application*

$$\begin{aligned} \phi: \mathbb{C} &\rightarrow \mathbb{C} \\ x &\mapsto \lambda x + \alpha \end{aligned} \quad (27.937)$$

avec  $\lambda \neq 0$  ( $\lambda \in \mathbb{C}$ ) et  $\alpha \in \mathbb{C}$ .

Soit  $q \in \mathbb{C}$ . Nous avons <sup>122</sup>

$$f_q(\phi(x)) = \lambda f_{\phi^{-1}(q)}(x). \quad (27.938)$$

*Démonstration.* En ce qui concerne la multiplication nous avons  $f_q(\lambda x) = \lambda x - q$ , et donc

$$\frac{1}{\lambda} f_q(\lambda x) = x - \frac{q}{\lambda} = f_{q/\lambda}(x). \quad (27.939)$$

Nous retenons :  $f_q(\lambda x) = \lambda f_{q/\lambda}(x)$ . En ce qui concerne la translation, c'est plus facile :

$$f_q(x + \alpha) = x + \alpha - q = f_{q-\alpha}(x). \quad (27.940)$$

Lorsque nous faisons les deux en même temps,

$$f_q(\phi(x)) = f_{q-\alpha}(\lambda x) = \lambda f_{(q-\alpha)/\lambda}(x). \quad (27.941)$$

Il suffit maintenant de remarquer que  $\phi^{-1}(q) = \frac{q-\alpha}{\lambda}$ , ce qui est facilement vérifié :

$$\phi\left(\frac{q-\alpha}{\lambda}\right) = \lambda\left(\frac{q-\alpha}{\lambda}\right) + \alpha = q. \quad (27.942)$$

□

LEMooBZUCooHWfolf

**Lemme 27.201** ([234]).

Soit un lacet de Jordan <sup>123</sup>  $\gamma: [a, b] \rightarrow \mathbb{C}$ . Nous notons  $\Gamma$  son image.

(1)  $\Gamma$  est compact.

(2) L'application <sup>124</sup>

$$\begin{aligned} I_\gamma: \mathcal{G}(\Gamma) &\rightarrow \mathbb{Z} \\ f &\mapsto \text{Ind}(f \circ \gamma, 0) \end{aligned} \quad (27.943)$$

est un morphisme surjectif de noyau  $\mathcal{E}(\Gamma)$ .

*Démonstration.* Nous notons  $\mathcal{L}(J, \mathbb{C}^*)$  le groupe multiplicatif des lacets  $J \rightarrow \mathbb{C}^*$ . Le fait que  $\Gamma$  soit compact est parce qu'il est l'image du compact  $[a, b]$  par l'application continue  $\gamma$  (théorème 7.274).

Nous considérons l'application

$$\begin{aligned} \xi: \mathcal{G}(\Gamma) &\rightarrow \mathcal{L}(J, \mathbb{C}^*) \\ f &\mapsto f \circ \gamma. \end{aligned} \quad (27.944)$$

(1)  $\xi$  est un morphisme Nous avons

$$\xi(fg) = (fg) \circ \gamma = (f \circ \gamma)(g \circ \gamma) = \xi(f)\xi(g). \quad (27.945)$$

122. L'application  $f_q$  est celle de la définition 27.199.

123. Définition 20.75.

124. Rappel définitions 27.197 pour  $\mathcal{G}(\Gamma)$  et  $\mathcal{E}(\Gamma)$ .

- (2)  **$\xi$  est injective** Nous nous souvenons de l'application quotient  $\tilde{\gamma}: S^1 \rightarrow \Gamma$ , qui est un homéomorphisme par le lemme 27.195(1). Nous nous souvenons également de l'application

$$\begin{aligned} q: [a, b] &\rightarrow \mathbb{C} \\ t &\mapsto \exp(2i\pi t/(b-a)) \end{aligned} \quad (27.946)$$

qui vérifie  $\gamma = \tilde{\gamma} \circ q$ . Nous montrons que ceci est l'inverse de  $\xi$  :

$$\begin{aligned} \theta: \mathcal{L}(J, \mathbb{C}^*) &\rightarrow \mathcal{G}(\Gamma) \\ \lambda &\mapsto \tilde{\lambda} \circ \tilde{\gamma}^{-1}. \end{aligned} \quad (27.947)$$

Soit  $\sigma \in \mathcal{L}(J, \mathbb{C}^*)$ . Nous avons

$$\xi(\theta(\sigma)) = \theta(\sigma) \circ \gamma = \tilde{\sigma} \circ \tilde{\gamma}^{-1} \circ \gamma = \tilde{\sigma} \circ \tilde{\gamma}^{-1} \circ \tilde{\gamma} \circ q = \tilde{\sigma} \circ q = \sigma. \quad (27.948)$$

- (3) **Conclusion 1** La proposition 26.56 donne un morphisme surjectif  $\psi: \mathcal{L}(J, \mathbb{C}^*) \rightarrow \mathbb{Z}$ . L'application  $I_\gamma$  dont nous parlons ici est exactement  $I_\gamma = \psi \circ \xi$ . Donc  $I_\gamma$  est un morphisme surjectif.

- (4) **Noyau** Nous devons encore montrer que le noyau de  $I_\gamma$  est  $\mathcal{E}(\Gamma)$ .

- (4a) **Dans un sens** Soit  $k \in \ker(I_\gamma)$ . Alors  $(\psi \circ \xi)(k) = 0$ , c'est-à-dire  $\psi(k \circ \gamma) = 0$ . La proposition 26.56(3) dit qu'alors  $k \circ \gamma$  est un lacet homotope dans  $\mathbb{C}^*$  à un lacet constant. Donc il existe une application continue  $H: [0, 1] \times [a, b] \rightarrow \mathbb{C}^*$  telle que

$$\begin{cases} H(0, t) = (k \circ \gamma)(t) \\ H(1, t) = 1 \end{cases} \quad \begin{aligned} (27.949a) \\ (27.949b) \end{aligned}$$

pour tout  $t$ . Vu que l'application  $\gamma: [a, b] \rightarrow \mathbb{C}^*$  est injective, nous pouvons considérer son inverse; pour la suite, quand nous écrirons  $\gamma^{-1}$ , c'est bien de  $\gamma^{-1}: \Gamma \rightarrow [a, b]$  que nous parlerons. C'est une application continue par le lemme 7.278. Nous pouvons donc considérer

$$\begin{aligned} L: [0, 1] \times \Gamma &\rightarrow \mathbb{C}^* \\ (u, z) &\mapsto H(u, \gamma^{-1}(z)). \end{aligned} \quad (27.950)$$

Cela est une application continue parce que  $H$  et  $\gamma$  le sont. Nous avons

$$L(0, z) = H(0, \gamma^{-1}(z)) = (k \circ \gamma)(\gamma^{-1}(z)) = k(z), \quad (27.951)$$

et

$$L(1, z) = 1. \quad (27.952)$$

Donc  $k$  est homotope à une application constante sur  $\Gamma$ . Le théorème de Borsuk 26.48 dit alors que  $k$  admet un logarithme continu sur  $\Gamma$ , c'est-à-dire que  $k \in \mathcal{E}(\Gamma)$ .

- (4b) **Dans l'autre sens** C'est le même raisonnement que le premier, mais un peu plus simple parce qu'il n'y a pas de débats sur la continuité. Soit  $k \in \mathcal{E}(\Gamma)$ . Le théorème de Borsuk indique que  $k$  est homotope dans  $\mathbb{C}^*$  à une application constante. Il existe une application continue  $H: [0, 1] \times \Gamma \rightarrow \mathbb{C}^*$  telle que

$$\begin{cases} H(0, z) = k(z) \\ H(1, z) = 1 \end{cases} \quad \begin{aligned} (27.953a) \\ (27.953b) \end{aligned}$$

pour tout  $z \in \Gamma$ . Le chemin  $k \circ \gamma$  est homotope à un lacet constant. En effet nous avons l'application continue

$$\begin{aligned} L: [0, 1] \times [a, b] &\rightarrow \mathbb{C}^* \\ (u, t) &\mapsto H(u, \gamma(t)) \end{aligned} \quad (27.954)$$

qui vérifie  $L(0, t) = (k \circ \gamma)(t)$  et  $L(1, t) = 1$ . Nous avons donc

$$0 = \psi(k \circ \gamma) = I_\gamma(k). \quad (27.955)$$

Nous avons prouvé que  $k \in \ker(I_\gamma)$ .

□

LEMooODIPooBZJPAW

**Lemme 27.202** ([234]).

Soit un lacet de Jordan  $\gamma$  dont l'image est  $\Gamma$ . Des fonctions  $f_0, f_1 \in \mathcal{G}(\Gamma)$  sont homotopes dans  $\mathbb{C}^*$  si et seulement si les lacets  $f_0 \circ \gamma$  et  $f_1 \circ \gamma$  le sont.

*Démonstration.* Dans les deux sens.

- (1)  $\Rightarrow$  Nous supposons que  $f_0$  et  $f_1$  sont homotopes : il existe une application continue  $H : [0, 1] \times \Gamma \rightarrow \mathbb{C}^*$  telle que

$$\begin{cases} H(0, z) = f_0(z) \\ H(1, z) = f_1(z) \end{cases} \quad (27.956a)$$

$$(27.956b)$$

pour tout  $z \in \Gamma$ . L'homotopie entre  $f_0 \circ \gamma$  et  $f_1 \circ \gamma$  est donnée par l'application continue

$$\begin{aligned} L : [0, 1] \times [a, b] &\rightarrow \mathbb{C}^* \\ (u, t) &\mapsto H(u, \gamma(t)). \end{aligned} \quad (27.957)$$

Elle vérifie  $L(0, t) = (f_0 \circ \gamma)(t)$  et  $L(1, t) = (f_1 \circ \gamma)(t)$ .

- (2)  $\Leftarrow$  Nous supposons l'existence d'une application continue  $H : [0, 1] \times [a, b] \rightarrow \mathbb{C}^*$  telle que  $H(0, t) = (f_0 \circ \gamma)(t)$  et  $H(1, t) = (f_1 \circ \gamma)(t)$ .

L'application  $\gamma : [a, b] \rightarrow \Gamma$  est continue et bijective. Sa réciproque  $\gamma^{-1}$  est donc également continue (lemme 7.278). Nous pouvons donc considérer l'application continue

$$\begin{aligned} L : [0, 1] \times \Gamma &\rightarrow \mathbb{C}^* \\ (u, z) &\mapsto H(u, \gamma^{-1}(z)). \end{aligned} \quad (27.958)$$

Elle vérifie

$$L(0, z) = H(0, \gamma^{-1}(z)) = (f_0 \circ \gamma)(\gamma^{-1}(z)) = f_0(z) \quad (27.959)$$

et  $L(1, z) = f_1(z)$ . Donc  $L$  est une homotopie dans  $\mathbb{C}^*$  entre  $f_0$  et  $f_1$ .

□

CORooDOMAooHtKPTe

**Corolaire 27.203** ([234]).

Si  $\Gamma$  est une courbe de Jordan, alors<sup>125</sup>  $G(\Gamma) \simeq \mathbb{Z}$ .

*Démonstration.* Nous allons utiliser le premier théorème d'isomorphisme 2.7. Si  $\gamma$  est un lacet de Jordan dont la courbe est  $\Gamma$ , le lemme 27.201 dit que  $I_\gamma : \mathcal{G}(\Gamma) \rightarrow \mathbb{Z}$  est un morphisme surjectif tel que  $\ker(I_\gamma) = \mathcal{E}(\Gamma)$ . Le théorème d'isomorphisme dit alors que

$$\frac{\mathcal{G}(\Gamma)}{\mathcal{E}(\Gamma)} = I_\gamma(\mathcal{G}(\Gamma)) = \mathbb{Z}. \quad (27.960)$$

□

Rappel, si  $K$  est compact dans  $\mathbb{C}$ , et si  $p \in \mathbb{C}$ , nous considérons la fonction  $f_p \in \mathcal{G}(K)$  par  $f_p(z) = z - p$ .

LEMooAKVFooFMHaOZ

**Lemme 27.204** ([1]).

Soient une courbe de Jordan  $\gamma$  ainsi que  $p \in \mathbb{C}$ . Nous avons

$$\text{Ind}(f_p \circ \gamma, 0) = \text{Ind}(\gamma, p). \quad (27.961)$$

PROPooUPPEooEQOOkh

**Proposition 27.205** ([234]).

Soit  $p \in \mathbb{C}$ . Si  $V$  est un disque ou un rectangle de centre  $p$ , alors  $\mathcal{G}(\partial V)$  est engendré par  $\mathcal{E}(\partial V)$  et par  $f_p$ .

<sup>125</sup>. Rappel :  $G(\Gamma)$  est défini en 27.197.

*Démonstration.* Ici  $\partial V$  peut être vu comme un chemin que nous allons noter aussi  $\gamma$ . Utilisant le lemme 27.204, nous avons

$$I_{\partial V}(f_p) = \text{Ind}(f_p \circ \partial V, 0) = \text{Ind}(\partial V, p). \quad (27.962)$$

(1) Si  $V$  est un disque Alors  $\partial V$  est le chemin

$$\begin{aligned} \gamma: [0, 2\pi] &\rightarrow \mathbb{C} \\ t &\mapsto p + Re^{it}. \end{aligned} \quad (27.963)$$

Nous avons alors le calcul suivant :

$$\text{Ind}(\gamma, p) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{\gamma'(t)}{\gamma(t) - p} dt \quad (27.964a)$$

$$= \frac{1}{2\pi i} \int_0^{2\pi} \frac{Re^{it}}{Re^{it}} dt \quad (27.964b)$$

$$= \frac{1}{2\pi i} \int_0^{2\pi} i dt \quad (27.964c)$$

$$= 1. \quad (27.964d)$$

(2) Si  $V$  est un rectangle Supposons que  $V$  ait une hauteur de  $h$  et une largeur de  $l$  tout en étant centré en  $p$ . Le chemin  $\partial V$  se décompose en quatre parties dont la première est

$$\begin{aligned} \gamma_1: [0, l] &\rightarrow \mathbb{C} \setminus \{p\} \\ t &\mapsto p - \frac{hi}{2} - \frac{l}{2} + t \end{aligned} \quad (27.965)$$

Je vous laisse voir les autres parties et calculer les intégrales<sup>126</sup>

Dans les deux cas nous avons  $\text{Ind}(\partial V, p) = 1$  et donc, en continuant (27.962),

$$I_{\partial V}(f_p) = 1. \quad (27.966)$$

Nous savons par le lemme 27.201 que

$$\begin{aligned} I_{\partial V}: \mathcal{G}(\partial V) &\rightarrow \mathbb{Z} \\ f &\mapsto \text{Ind}(f \circ \gamma, 0) \end{aligned} \quad (27.967)$$

est un morphisme surjectif de noyau  $\mathcal{E}(\Gamma)$ . Le fait que ce soit un morphisme implique que

$$I_{\partial V}(f_p^n) = nI_{\partial V}(f_p) = n. \quad (27.968)$$

Mais  $G(\partial V) = \mathcal{G}(\partial V)/\mathcal{E}(\partial V)$ . Donc nous avons déjà prouvé que  $\{f_p^k\}_{k \in \mathbb{N}}$  contient un élément de chaque classe dans  $\mathcal{G}(\partial V)/\mathcal{E}(\partial V)$ .

Soit  $f \in \mathcal{G}(\partial V)$ . En notant  $[.]$  la classe par rapport à  $\mathcal{E}(\partial V)$ , il existe  $k \in \mathbb{N}$  tel que

$$f \in [I_{\partial V}(f_p^k)]. \quad (27.969)$$

Autrement dit, en posant  $\alpha = I_{\partial V}(f_p^k)$ , il existe  $h \in \mathcal{E}(\partial V)$  tel que  $f = \alpha h$ . Nous avons donc bien prouvé que  $\{f_p, \mathcal{E}(\partial V)\}$  engendrent  $\mathcal{G}(\partial V)$ .  $\square$

**Proposition 27.206** ([234]).

Soit un compact  $K$  de  $\mathbb{C}$  ainsi que  $p, q$  deux points dans la même composante connexe de  $\mathbb{C} \setminus K$ . Alors

(1)  $f_p/f_q \in \mathcal{E}(K)$ .

126. Je n'ai pas vérifié ; si ça pose un problème, écrivez-moi.

ITEMooUJLPooApgXIF

(2) Si  $p$  est dans la composante non bornée<sup>127</sup> de  $\mathbb{C} \setminus K$ , alors  $f_p \in \mathcal{E}(K)$ .

*Démonstration.* Soient  $p$  et  $q$  dans la même composante connexe de  $\mathbb{C} \setminus K$ . La proposition 9.68 nous autorise à considérer un chemin  $\gamma: [0, 1] \rightarrow \mathbb{C} \setminus K$  tel que  $\gamma(0) = p$  et  $\gamma(1) = q$ . Nous posons

$$\begin{aligned} F: [0, 1] \times K &\rightarrow \mathbb{C}^* \\ (t, z) &\mapsto f_{\gamma(t)}(z) = z - \gamma(t). \end{aligned} \quad (27.970)$$

C'est une application continue vérifiant  $F(0, z) = f_p(z)$  et  $F(1, z) = f_q(z)$ . Donc les applications  $f_p$  et  $f_q$  sont homotopes dans  $\mathbb{C}^*$ . La proposition 26.16 nous indique qu'alors l'application  $f_p/f_q$  est homotope à 1. Le théorème de Borsuk 26.48 conclut que  $f_p/f_q \in \mathcal{E}(K)$ .

Soit  $p$  dans la composante non bornée de  $\mathbb{C} \setminus K$ . Vu que  $K$  est borné, nous pouvons prendre  $r > 0$  tel que  $K \subset B(0, r)$ . Soit  $p_0 \notin \overline{B(0, r)}$ ; ce point est également dans la composante non bornée de  $K$ . Par le point (1), nous avons  $f_p/f_{p_0} \in \mathcal{E}(K)$ .

Nous allons prouver que  $f_{p_0} \in \mathcal{E}(K)$ . Pour cela nous posons

$$\begin{aligned} F: [0, 1] \times K &\rightarrow \mathbb{C}^* \\ (t, z) &\mapsto tz - p_0. \end{aligned} \quad (27.971)$$

Vu que  $t \in [0, 1]$  et que  $z \in K \subset B(0, r)$ , nous avons toujours  $tz \in B(0, r)$ . Étant donné que  $p_0$  n'est pas dans  $B(0, r)$ , l'application  $F$  prend bien ses valeurs dans  $\mathbb{C}^*$ .

L'application  $F$  vérifie également  $F(0, z) = -p_0$  et  $F(1, z) = z - p_0 = f_{p_0}(z)$ . Donc  $f_{p_0}$  est homotope dans  $\mathbb{C}^*$  à l'application constante  $-p_0$ . Nous en déduisons que  $f_{p_0} \in \mathcal{E}(K)$ . Au final,

$$f_p = f_{p_0} \frac{f_p}{f_{p_0}} \in \mathcal{E}(K) \quad (27.972)$$

parce que  $f_{p_0}$  et  $f_p/f_{p_0}$  sont dans  $\mathcal{E}(K)$ . □

LEMooOTMGooZeUoCy

**Lemme 27.207** ([234]).

Soient des composantes connexes bornées  $\mathcal{O}_1, \dots, \mathcal{O}_N$  de  $\mathbb{C} \setminus K$  que nous supposons deux à deux disjointes. Soient  $p_i \in \mathcal{O}_i$ . Si  $n_1, \dots, n_N \in \mathbb{Z}$  sont des entiers non tous nuls, alors

$$\prod_{i=1}^N f_{p_i}^{n_i} \notin \mathcal{E}(K). \quad (27.973)$$

*Démonstration.* En plusieurs points.

- (1) **Par l'absurde** Nous notons  $f = \prod_{i=1}^N f_{p_i}^{n_i}$  et nous supposons par l'absurde que  $f \in \mathcal{E}(K)$ . Par le théorème de Borsuk 26.48, nous avons une extension continue  $\tilde{f}: \mathbb{C} \rightarrow \mathbb{C}^*$ .
- (2)  $\partial \mathcal{O}_i \subset K$  C'est le lemme 7.250.
- (3)  **$p_j$  n'est pas dans  $\bar{\mathcal{O}}_i$**  Nous savons que  $p_j \in \mathcal{O}_j$ . Mais  $\partial \mathcal{O}_i \subset K$  et  $K \cap \mathcal{O}_j = \emptyset$ . Donc  $p_j$  n'est pas dans  $\partial \mathcal{O}_i$ . Comme les  $\mathcal{O}_i$  sont disjointes, nous avons aussi  $p_j \notin \mathcal{O}_i$ . Bref,

$$p_j \notin \bar{\mathcal{O}}_i. \quad (27.974)$$

- (4) **Quelques fonctions** Nous posons

$$\begin{aligned} G_i: \bar{\mathcal{O}}_i &\rightarrow \mathbb{C}^* \\ z &\mapsto \prod_{j \neq i} (z - p_j)^{n_j}. \end{aligned} \quad (27.975)$$

Le fait que  $G_i$  prenne ses valeurs dans  $\mathbb{C}^*$  est parce que  $p_j \notin \bar{\mathcal{O}}_i$  pour tout  $i \neq j$ .

127. Il y en a une seule, c'est le lemme 26.50.

Nous reprenons notre extension continue  $\tilde{f}: \mathbb{C} \rightarrow \mathbb{C}^*$  et nous posons

$$\begin{aligned} F_i: \bar{\mathcal{O}}_i &\rightarrow \mathbb{C}^* \\ z &\mapsto \frac{\tilde{f}(z)}{G_i(z)}. \end{aligned} \quad (27.976)$$

Vu que  $G_i$  ne s'annule pas sur  $\bar{\mathcal{O}}_i$ , cette fonction  $F_i$  est bien définie et continue.

Nous nous fixons  $i$ , et nous allons prouver que  $n_i = 0$ . Nous posons  $D = B(0, R)$  avec  $R$  assez grand pour contenir  $\bar{\mathcal{O}}_i$ . Posons

$$\begin{aligned} H_i: \overline{B(0, R)} &\rightarrow \mathbb{C}^* \\ z &\mapsto \begin{cases} F_i(z) & \text{si } z \in \bar{\mathcal{O}}_i \\ (z - p_i)^{n_i} & \text{sinon.} \end{cases} \end{aligned} \quad (27.977)$$

- (5)  **$H_i$  est continue** L'application  $F_i$  est continue sur  $\bar{\mathcal{O}}_i$ . L'application  $z \mapsto (z - p_i)^{n_i}$  est continue sur  $\mathbb{C}$ .

Soit  $z \in \partial\mathcal{O}_i$ . Nous avons  $z \in \partial\mathcal{O}_i \subset K$  et donc  $\tilde{f}(z) = f(z)$ . Donc

$$F_i(z) = \frac{f(z)}{G_i(z)} = f_{p_i}(z)^{n_i}. \quad (27.978)$$

Les deux applications  $F_i$  et  $z \mapsto (z - p_i)^{n_i}$  sont égales sur le bord  $\partial\mathcal{O}_i$ . Donc  $H_i$  est continue sur  $\overline{B(0, R)}$ .

- (6) **Logarithme continu** La partie  $\overline{B(0, R)}$  est convexe et compacte. Le corolaire 26.49 dit que  $H_i$  admet un logarithme continu. La proposition 26.59 nous affirme que pour tout lacet  $\gamma: [0, 1] \rightarrow \overline{B(0, R)}$ , nous avons  $\text{Ind}(H_i \circ \gamma, 0) = 0$ . En particulier le chemin  $\partial B(0, R)$  est dans  $\overline{B(0, R)}$ . Avant de calculer  $\text{Ind}(H_i \circ \gamma, 0)$ , remarquons que  $\gamma(t) \in \overline{B(0, R)}$  et reste donc en-dehors de  $\bar{\mathcal{O}}_i$ . Donc

$$H_i(\gamma(t)) = (\gamma(t) - p_i)^{n_i} \quad (27.979)$$

et

$$(H_i \circ \gamma)'(t) = n_i \gamma'(t) (\gamma(t) - p_i)^{n_i - 1}. \quad (27.980)$$

De plus si vous voulez une forme explicite de  $\gamma$ , c'est pas compliqué : c'est un cercle de rayon  $R$ . Nous avons un calcul :

$$0 = \text{Ind}(H_i \circ \gamma, 0) = \frac{1}{2\pi i} \int_{H_i \circ \gamma} \frac{dz}{z} \quad (27.981a)$$

$$= \frac{1}{2\pi i} \int_0^1 \frac{(H_i \circ \gamma)'(t)}{(H_i \circ \gamma)(t)} dt \quad (27.981b)$$

$$= \frac{1}{2\pi i} \int_0^1 \frac{n_i \gamma'(t) (\gamma(t) - p_i)^{n_i - 1}}{(\gamma(t) - p_i)^{n_i}} dt \quad (27.981c)$$

$$= \frac{1}{2\pi i} \int_0^1 \frac{n_i \gamma'(t)}{\gamma(t) - p_i} dt \quad (27.981d)$$

$$= n_i \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z - p_i} \quad (27.981e)$$

$$= n_i \text{Ind}(\gamma, p_i). \quad (27.981f)$$

Vu que  $p_i \in B(0, R)$  et que  $\gamma$  est un cercle autour, nous avons  $\text{Ind}(\gamma, p_i) = 1$ . Donc

$$0 = \text{Ind}(H_i \circ \gamma, 0) = n_i. \quad (27.982)$$

□

LEMooEJRMooNjHMOV

**Lemme 27.208** ([234, 1]).

Nous notons  $(\mathcal{O}_i)_{i \in I}$  les composantes connexes bornées de  $\mathbb{C} \setminus K$ . Pour chaque  $i \in I$  nous considérons  $p_i \in \mathcal{O}_i$ . Alors

$$\mathcal{G}(K) = \text{gr}(\mathcal{E}(K), \{f_{p_i}\}_{i \in I}). \quad (27.983)$$

*Démonstration.* Nous commençons par prouver que si  $q \in \mathbb{C} \setminus K$ , alors

$$f_q \in \text{gr}(\mathcal{E}(K), \{f_{p_i}\}_{i \in I}). \quad (27.984)$$

Il y a deux possibilités : soit  $q$  est dans la composante non bornée, soit il est dans un des  $\mathcal{O}_i$ .

(1) **Si  $q$  est dans la composante non bornée** Alors la proposition 27.206(2) dit que  $f_q \in \mathcal{E}(K)$ .

(2) **Si  $q \in \mathcal{O}_i$**  Alors la proposition 27.206(1) dit que  $f_q/f_{p_i} \in \mathcal{E}(K)$ . Dans ce cas nous avons

$$f_q = f_{p_i} \frac{f_q}{f_{p_i}} \in \text{gr}(f_{p_i}, \mathcal{E}(K)). \quad (27.985)$$

Soit  $f \in \mathcal{G}(K)$ . Nous devons prouver l'existence de  $g \in \mathcal{E}(K)$ ,  $q_1, \dots, q_l \in \mathbb{C} \setminus K$  et  $m_1, \dots, m_l \in \mathbb{Z}$  tels que

$$f = g \prod_{i=1}^l f_{q_i}^{m_i}. \quad (27.986)$$

Nous commençons par supposer que  $K \subset Q = [0, 1] \times [0, 1]$ .

(1) **Les petits carrés** Soit  $n \in \mathbb{N}$ . Nous subdivisons le carré  $Q$  en carrés fermés de taille  $1/n$  ; il y a  $n^2$  tels carrés. Nous notons  $Q_1, \dots, Q_l$  ceux qui n'ont pas d'intersection avec  $K$ , et nous notons  $A$  l'union des autres.

Grâce au lemme 10.35, nous savons que pour tout  $z \in A$ , nous avons  $d(z, K) \leq \sqrt{2}/n$ .

(2) **Prolongement** La partie  $K$  est fermée,  $f$  est continue dessus. Le théorème de Tietze 27.176 donne un prolongement continu  $F: \mathbb{C} \rightarrow \mathbb{C}$ . Nous avons

$$K \subset S = \{z \in \mathbb{C} \text{ tel que } F(z) \neq 0\}. \quad (27.987)$$

Vu que  $F$  est continue,  $S$  est ouverte.

(3) **Grand  $n$**  Nous prouvons que  $n$  peut être choisi assez grand pour que  $F$  ne s'annule pas sur  $A$ . La fonction  $F$  est continue et ne s'annule pas sur l'ouvert  $S$  qui contient  $K$ . Si  $S = \mathbb{C}$ , alors n'importe quel  $n$  convient.

Sinon nous considérons

$$\begin{aligned} \varphi: K &\rightarrow [0, \infty[ \\ z &\mapsto d(z, S^c). \end{aligned} \quad (27.988)$$

L'application  $\varphi$  est continue sur le compact  $K$ . Elle a donc un minimum global que nous nommons  $s$ . Pour tout  $z \in K$  nous avons  $d(z, S^c) \geq s$ . Donc pour tout  $z \in K$  nous avons  $B(z, s/2) \subset S$ . Nous prenons  $n$  assez grand pour avoir

$$\frac{\sqrt{2}}{n} < \frac{s}{2}. \quad (27.989)$$

EQooEWQYooWYsImg

Avec ça nous avons  $A \subset S$ . En effet, soit  $z \in A$ . Nous avons  $d(z, K) < \sqrt{2}/n < s/2$ . Donc il existe  $k \in K$  tel que

$$z \in B(k, \frac{s}{2}). \quad (27.990)$$

Comme  $k \in K$ , nous avons  $d(k, S^c) = \varphi(k) \geq s$  et donc  $B(k, s/2) \subset S$ . Nous avons prouvé que  $z \in S$ .

Bref, nous prenons  $n$  assez grand pour vérifier (27.989), et nous avons alors

$$A \subset \{z \in \mathbb{C} \text{ tel que } F(z) \neq 0\}. \quad (27.991)$$

- (4) **Une propriété à prouver** SPooVQHTooUNukEE Nous allons prouver l'énoncé suivant. Soit  $k = 0, \dots, l$ , soit une application continue  $h: K \rightarrow \mathbb{C}^*$  admettant un prolongement continu  $F: A \cup (Q_1 \cup \dots \cup Q_k) \rightarrow \mathbb{C}^*$ . Alors il existe  $m \in \mathbb{N}$  tel que l'application  $h/f_{q_{k+1}}^m$  admette un prolongement continu  $\tilde{F}: A \cup (Q_1 \cup \dots \cup Q_{k+1}) \rightarrow \mathbb{C}^*$ .
- (5) **Note pour  $f$**   
Jusqu'ici, nous avons prouvé que  $f$  vérifie les hypothèses de la récurrence pour  $k = 0$ . De plus, cette propriété à prouver n'est pas à prouver par récurrence sur  $k$ . Le  $k$  est fixé, et nous prouvons la propriété de façon directe. Une récurrence va venir un peu plus loin.
- (6) **Preuve de la propriété** Soit  $h: K \rightarrow \mathbb{C}^*$  continue et admettant un prolongement continu  $F: A \cup (Q_1 \cup \dots \cup Q_k) \rightarrow \mathbb{C}^*$ . Notons  $B_k = A \cup (Q_1 \cup \dots \cup Q_k)$  et posons

$$g: \partial Q_{k+1} \rightarrow \mathbb{C}^*$$

$$z \mapsto \begin{cases} F(z) & \text{si } z \in Q_{k+1} \cap B_k \\ s(z) & \text{sinon.} \end{cases} \quad (27.992)$$

Voyons ce que  $s$  peut être pour que  $g$  soit continue. La partie  $\partial Q_{k+1}$  est le bord d'un carré et est donc constitué de 4 côtés. La valeur de  $g$  sur certains de ces côtés (ceux qui sont aussi dans  $B_k$ ) est prescrite par  $F$ . Sur les autres côtés, nous pouvons mettre n'importe quoi pourvu que ce soit égal à  $F$  sur les coins et que  $s(z) \neq 0$ . Étant donné que les valeurs de  $F$  sur les coins est non nulle et que  $\mathbb{C}^*$  est connexe par arcs, c'est possible.

Nous notons  $q_{k+1}$  le centre de  $Q_{k+1}$ . La proposition 27.205 s'applique au compact  $\partial Q_{k+1}$ .

$$g \in \mathcal{G}(\partial Q_{k+1}) = \text{gr}(\mathcal{E}(\partial Q_{k+1}), f_{q_{k+1}}). \quad (27.993)$$

Il existe donc  $\xi \in \mathcal{E}(\partial Q_{k+1})$  et  $m \in \mathbb{Z}$  tels que

$$g(z) = \xi(z) f_{q_{k+1}}^m(z). \quad (27.994)$$

Prouvons que  $Q_{k+1} \cap B_k$  est une union de côtés de  $Q_{k+1}$ . Nous avons

$$Q_{k+1} \cap B_k = (Q_{k+1} \cap A) \cup (Q_{k+1} \cap Q_1) \cup \dots \cup (Q_{k+1} \cap Q_k). \quad (27.995) \quad \text{EQooSPSQooVFrzCI}$$

Nous nous souvenons que  $A = \bigcup_{i=1}^s R_i$  où  $R_i$  sont les carrés qui intersectent  $K$ . Donc  $R_i \cap Q_j$  est au maximum les côtés de  $Q_j$ . Bref, chacun des termes de l'union à droite de (27.995) est une union de côtés de  $Q_{k+1}$ .

L'application  $\xi \in \mathcal{E}(\partial Q_{k+1})$  admet un logarithme continu et admet donc une extension continue  $\mathbb{C} \rightarrow \mathbb{C}^*$  (théorème de Borsuk 26.48). Nous considérons cette extension seulement sur  $Q_{k+1}$  et nous l'appelons  $G: Q_{k+1} \rightarrow \mathbb{C}^*$ . Nous posons

$$\tilde{F}: A \cup (Q_1 \cup \dots \cup Q_{k+1}) \rightarrow \mathbb{C}^*$$

$$z \mapsto \begin{cases} G(z) & \text{si } z \in Q_{k+1} \\ \frac{F(z)}{(z - q_{k+1})^m} & \text{si } z \in A \cup (Q_1 \cup \dots \cup Q_k). \end{cases} \quad (27.996)$$

Voyons que cette fonction est continue en montrant que les deux morceaux ont les mêmes valeurs sur l'intersection. Soit  $z \in Q_{k+1} \cap (A \cup Q_1 \cup \dots \cup Q_k)$ . Alors  $z \in \partial Q_{k+1}$  parce que nous avons dit que (27.995) est dans le bord dans  $Q_{k+1}$ . Dans ce cas nous avons

$$G(z) = \xi(z) = \frac{g(z)}{f_{q_{k+1}}^m(z)} \quad (27.997a)$$

$$= \frac{F(z)}{f_{q_{k+1}}^m(z)}, \quad (27.997b)$$

$$(27.997c)$$



et donc bien la continuité de  $\tilde{F}$ .

Si  $z \in K$ , étant donné que  $K \subset A$  nous avons

$$\tilde{F}(z) = \frac{F(z)}{(z - q_{k+1})^m} = \frac{h(z)}{f_{q_{k+1}}^m(z)}, \quad (27.998)$$

donc  $\tilde{F}$  prolonge bien  $h/f_{q_{k+1}}^m$ .

- (7) **Récurrence** En appliquant la propriété (4)  $l$  fois, nous construisons une suite d'entiers  $(m_i)_{i=1, \dots, l}$  et une application continue  $F_l: A \cup (Q_1 \cup \dots \cup Q_l) = Q \rightarrow \mathbb{C}^*$  qui prolonge  $f / \prod_{i=1}^l f_{q_i}^{m_i}$ .

Sur  $K$  nous avons alors

$$f = F_l \prod_{i=1}^l f_{q_i}^{m_i}. \quad (27.999)$$

C'est une application continue sur le convexe  $Q$ . Donc  $F_l \in \mathcal{E}(Q)$  et donc  $F_l \in \mathcal{E}(K)$ .

Maintenant vous croyez avoir fini mais non. Vous devez encore vérifier que tout est encore valable si  $K$  n'est pas contenu dans  $[0, 1] \times [0, 1]$ .

Soit un compact général  $K \subset \mathbb{C}$ . Pour  $\lambda > 0$  et  $\alpha \in \mathbb{C}$ , nous posons  $\phi(x) = \lambda x + \alpha$ . Il existe forcément un choix de  $\lambda$  et de  $\alpha$  tels que  $\phi(K) \subset [0, 1] \times i[0, 1]$ .

Soit  $f \in \mathcal{G}(K)$ . Nous appliquons le résultat déjà prouvé à l'application  $f \circ \phi^{-1} \in \mathcal{G}(\phi(K))$ . Nous avons  $g \in \mathcal{E}(\phi(K))$ ,  $q_1, \dots, q_l \in \mathbb{C} \setminus \phi(K)$  et  $m_1, \dots, m_l \in \mathbb{Z}$  tels que, pour tout  $y \in \phi(K)$ ,

$$(f \circ \phi^{-1})(y) = g(y) \prod_{i=1}^l f_{q_i}^{m_i}(y). \quad (27.1000)$$

Pour  $x \in K$  nous avons donc

$$f(x) = (g \circ \phi)(x) \prod_{i=1}^l f_{q_i}^{m_i}(\phi(x)). \quad (27.1001)$$

C'est le moment de sortir le lemme 27.200 :

$$f(x) = (g \circ \phi)(x) \prod_{i=1}^l f_{q_i}^{m_i}(\phi(x)) = (g \circ \phi)(x) \prod_{i=1}^l \lambda^{m_i} f_{\phi^{-1}(q_i)}^{m_i}(x). \quad (27.1002)$$

En posant

$$\begin{cases} \tilde{g}(x) = \left( \prod_{i=1}^l \lambda^{m_i} \right) (g \circ \phi)(x) & (27.1003a) \\ \tilde{m}_i = m_i & (27.1003b) \\ \tilde{q}_i = \phi^{-1}(q_i), & (27.1003c) \end{cases}$$

nous avons  $\tilde{q}_i \in \mathbb{C} \setminus K$ ,  $\tilde{g} \in \mathcal{E}(K)$  et  $f = \tilde{g} \prod_{i=1}^l f_{\tilde{q}_i}$ . □

LEMooBHOGooXxBYGA

### Lemme 27.209.

Une partie ouverte de  $\mathbb{C}$  a au plus un nombre dénombrable de composantes connexes.

*Démonstration.* Soit un ouvert  $A$  de  $\mathbb{C}$ . Nous notons  $\mathcal{B}$  une base dénombrable d'ouverts de  $\mathbb{C}$  (proposition 7.188). Une composante connexe de  $A$  est ouverte<sup>128</sup>. Les composantes connexes de  $A$  sont disjointes, et chacune est l'union d'une quantité (au plus) dénombrable d'éléments de  $\mathcal{B}$ . Si il y avait une quantité non dénombrable de composantes connexes, cela ferait une quantité non dénombrable d'éléments différents de  $\mathcal{B}$ . Donc non. □

<sup>128</sup>. Proposition 7.249 et le fait que  $\mathbb{C}$  est localement convexe.

**Théorème 27.210.**

Soit un compact  $K$  dans  $\mathbb{C}$ . Alors

- (1) Le nombre de composantes connexes bornées de  $\mathbb{C} \setminus K$  est fini ou dénombrable.  
 (2) En notant  $N$  le nombre de composantes connexes bornées de  $\mathbb{C} \setminus K$ , nous avons

$$G(K) = \mathcal{G}(K)/\mathcal{E}(K) \simeq \mathbb{Z}^N \quad (27.1004)$$

où, si  $N$  est infini,  $\mathbb{Z}^N$  désigne l'ensemble des suites finies dans  $\mathbb{Z}$ .

*Démonstration.* Le fait que le nombre de composantes connexes bornées de  $\mathbb{C} \setminus K$  est au maximum dénombrable est le lemme 27.209.

Soient  $\{\mathcal{O}_i\}_{i=0,\dots,N}$  les composantes connexe bornées de  $\mathbb{C} \setminus K$  (avec  $i \in \mathbb{N}$  dans le cas où  $N$  il y a une infinité dénombrable de composants connexes). Pour chaque  $i$  nous prenons un élément  $p_i \in \mathcal{O}_i$ . Nous notons aussi  $\pi: \mathcal{G}(K) \rightarrow G(K)$  le morphisme canonique. Nous posons

$$\begin{aligned} \phi: \mathbb{Z}^N &\rightarrow G(K) \\ n &\mapsto \prod_{i=1}^N \pi(f_{p_i})^{n_i}. \end{aligned} \quad (27.1005)$$

Notez que le produit a toujours un sens parce que  $n \in \mathbb{Z}^N$  est toujours une suite finie (de taille  $N$  ou arbitraire) d'entiers. Pas de soucis de convergence.

- (1)  $\phi$  est un morphisme Sur  $\mathbb{Z}^n$ , nous mettons l'addition  $(a+b)_i = a_i + b_i$ . Si  $N = \infty$  et si  $a$  et  $b$  ont des longueurs différentes, on prend pour  $N$  le maximum des deux longueurs. Nous avons :

$$\phi(a)\phi(b) = \prod_{i=1}^N \pi(f_{p_i})^{a_i} \prod_{j=1}^N \pi(f_{p_j})^{b_j} \quad (27.1006a)$$

$$= \prod_{i=1}^N \pi(f_{p_i})^{a_i+b_i} \quad (27.1006b)$$

$$= \phi(a+b). \quad (27.1006c)$$

- (2)  $\phi$  est injective Soit  $a \in \mathbb{Z}^N$ . Si  $a \neq 0$ , alors le lemme 27.207 nous dit que  $\prod_{i=1}^N f_{p_i}^{a_i} \notin \mathcal{E}(K)$ . Vu que dans  $G(K)$ , le zéro est  $\mathcal{E}(K)$ , nous avons  $\prod_{i=1}^N \pi(f_{p_i})^{a_i} \neq 0$  dès qu'un des  $a_i$  est non nul.

- (3)  $\phi$  est surjective Soit  $F \in G(K)$ . Vu que  $\pi$  est surjective, il existe  $f \in \mathcal{G}(K)$  tel que  $\pi(f) = F$ . Le lemme 27.208 nous dit qu'il existe  $g \in \mathcal{E}(K)$  et  $a \in \mathbb{Z}^N$  tels que

$$f = g \prod_{i=1}^N f_{p_i}^{a_i}. \quad (27.1007)$$

Nous avons alors

$$F = \pi(f) = \pi(g)\pi\left(\prod_{i=1}^N f_{p_i}^{a_i}\right) = \prod_{i=1}^N \pi(f_{p_i})^{a_i} = \phi(a). \quad (27.1008)$$

□

**Définition 27.211** ([234, 1]).

Si  $(G, +)$  est un groupe abélien, la **pseudo-dimension** de  $G$  est le cardinal maximum des familles  $\mathbb{Z}$ -indépendantes de  $G$ .

Dans ce contexte, une famille  $\mathbb{Z}$ -indépendante est une partie  $\{a_i\}_{i \in I}$  de  $G$  telle que si  $\sum_i z_i a_i = 0$  avec  $z_i \in \mathbb{Z}$ , alors  $z_i = 0$  pour tout  $i$ .

Si il existe une partie libre idempotente à un ensemble  $A$ , mais pas de parties libres surpotentes<sup>129</sup> à  $A$ , nous disons que la pseudo-dimension de  $G$  est « la cardinalité de  $A$  ».

Dans tous les cas, nous notons  $\text{pdim}(G)$  la pseudo-dimension de  $G$ .

LEMooPASKooFGJgBz

**Lemme 27.212** ([234, 1]).

À propos de pseudo-dimension.

- (1) La pseudo-dimension de  $\mathbb{Z}^n$  est  $n$ .
- (2) La pseudo-dimension de  $\mathbb{Z}^\infty$ <sup>130</sup> est la cardinalité de  $\mathbb{N}$ .

*Démonstration.* Nous commençons par le cas de  $\mathbb{Z}^n$  et nous ferons  $\mathbb{Z}^\infty$  après.

- (1)  $\geq n$  Prenons la base canonique  $(e_i)_{i=1,\dots,n}$  de  $\mathbb{Z}^n$ . C'est une partie  $\mathbb{Z}$ -indépendante de  $\mathbb{Z}^n$ . Donc la pseudo-dimension de  $\mathbb{Z}^n$  est plus grande ou égale à  $n$ .
- (2)  $\leq n$  Dans l'autre sens, supposons que  $\{a_i\}_{i \in I}$  soit  $\mathbb{Z}$ -indépendant dans  $\mathbb{Z}^n$ . Pour rappel, la définition 4.1 de partie libre ne parle que de sommes finies.

Prouvons que  $\{a_i\}_{i \in I}$  est  $\mathbb{Q}$ -libre dans  $\mathbb{Q}^n$ . Soient une partie finie  $J$  de  $I$  ainsi que des rationnels  $\{q_j\}_{j \in J}$  que  $\sum_{j \in J} q_j a_j = 0$ . Nous pouvons utiliser le lemme 1.501 pour considérer  $k \in \mathbb{N} \setminus \{0\}$  tel que  $kq_j \in \mathbb{Z}$  pour tout  $j$  (ici il est crucial que  $J$  soit fini, sinon on ne peut pas prendre un maximum). Dans ce cas nous avons  $\sum_{j \in J} kq_j a_j = 0$ , ce qui montre que  $\{a_i\}$  n'est pas  $\mathbb{Z}$ -indépendante.

L'intérêt de  $\mathbb{Q}$  est d'être un corps (alors que  $\mathbb{Z}$  ne l'est pas). Nous savons que  $\mathbb{Q}^n$  est un  $\mathbb{Q}$ -espace vectoriel de dimension  $n$ . Il ne contient donc pas de parties  $\mathbb{Q}$ -libres de cardinal plus grand que  $n$ .

Nous passons maintenant à la preuve pour  $\mathbb{Z}^\infty$ [1]. L'existence d'une partie dénombrable et  $\mathbb{Z}$ -libre dans  $\mathbb{Z}^\infty$  est facile : il faut juste prendre  $e_i$  étant la suite nulle partout sauf en  $i$ . La partie  $\{e_i\}_{i \in \mathbb{N}}$  est  $\mathbb{Z}$ -libre dans  $\mathbb{Z}^\infty$ .

Nous devons maintenant montrer que toute partie libre de  $\mathbb{Z}^\infty$  est au maximum dénombrable. Soit une telle partie  $A$ . En posant

$$A_n = \{x \in A \text{ tel que } x_n \neq 0 \text{ et } x_k = 0 \text{ si } k > n\}. \quad (27.1009)$$

Étant donné que  $A$  ne contient que des suites finies,  $A = \bigcup_{n \in \mathbb{N}} A_n$ .

D'abord  $A_n$  est  $\mathbb{Z}$ -libre dans  $\mathbb{Z}^n$ . En effet considérons une partie finie  $\{a_i\}_{i \in I}$  de  $A_n$  ainsi que des entiers  $\{z_i\}_{i \in I}$  tels que  $\sum_{i \in I} z_i a_i = 0$ . Vu que  $\{a_i\}_{i \in I}$  est également une partie finie de  $A$ , nous en déduisons que  $z_i = 0$  pour tout  $i$ .

La partie  $A_n$  étant libre dans  $\mathbb{Z}^n$ , elle est finie (de cardinal au maximum  $n$ ). La partie  $A$  étant une union dénombrable de parties finies, elle est au maximum dénombrable.  $\square$

LEMooEYDNooEUKUpn

**Lemme 27.213.**

Deux groupes abéliens isomorphes ont même pseudo-dimension.

*Démonstration.* Soient des groupes abéliens  $G$  et  $H$  ainsi qu'un isomorphisme  $\phi: G \rightarrow H$ . Soit une partie  $\mathbb{Z}$ -indépendante  $\{a_i\}_{i \in I}$  dans  $G$ . Nous montrons que  $\{\phi(a_i)\}_{i \in I}$  est  $\mathbb{Z}$ -indépendante dans  $H$ .

En effet si  $\sum_{i \in I} z_i \phi(a_i) = 0$ , alors  $0 = \sum_i \phi(z_i a_i) = \phi(\sum_i z_i a_i)$ , et donc  $\sum_i z_i a_i = 0$ , ce qui entraîne  $z_i = 0$ .

Cela prouve que  $\text{pdim}(H) \geq \text{pdim}(G)$ . En inversant les rôles de  $G$  et  $H$  (et en utilisant  $\phi^{-1}$ ), nous prouvons que  $\text{pdim}(G) \geq \text{pdim}(H)$ .

Si une des pseudo-dimensions est infinie, le raisonnement donné doit être récrit en disant que  $G$  possède une partie  $\mathbb{Z}$ -libre de telle cardinalité si et seulement si  $H$  en possède une.  $\square$

CORooQNUllooKltWVD

**Corolaire 27.214** ([234]).

Soient des compacts  $K_1$  et  $K_2$  homéomorphes dans  $\mathbb{C}$ . Alors les parties  $\mathbb{C} \setminus K_1$  et  $\mathbb{C} \setminus K_2$  ont le même nombre de composantes connexes.

129. Définition 1.65.

130. Pour rappel,  $\mathbb{Z}^\infty$  est l'ensemble des suites finies d'éléments de  $\mathbb{Z}$ .

*Démonstration.* Vu que  $K_1$  est homéomorphe à  $K_2$ , les groupes  $G(K_1)$  et  $G(K_2)$  sont isomorphes par le lemme 27.198. Le théorème 27.210(2) nous enseigne que, si on désigne par  $N_i$  le nombre de composantes connexes de  $\mathbb{C} \setminus K_i$ , alors  $G(K_i) = \mathbb{Z}^{N_i}$ . Le fait que  $G(K_1)$  soit isomorphe à  $G(K_2)$  implique

$$\mathbb{Z}^{N_1} \simeq \mathbb{Z}^{N_2}. \quad (27.1010)$$

Le lemme 27.213 indique alors que  $\text{pdim}(\mathbb{Z}^{N_1}) = \text{pdim}(\mathbb{Z}^{N_2})$ . Et enfin le lemme 27.212 conclut que  $N_1 = N_2$ .  $\square$

Dans le lemme suivant,  $S^1(a, r)$  est la sphère centrée en  $a$  et de rayon  $r$  :  $S^1(a, r) = \{z \in \mathbb{C} \text{ tel que } |a - z| = r\}$ .

LEMooUSJZooFVAITa

### Lemme 27.215.

La partie  $\mathbb{C} \setminus S^1(0, 1)$  a exactement 2 composantes connexes.

*Démonstration.* Posons  $A = \{z \in \mathbb{C} \text{ tel que } |z| < 1\}$  et  $B = \{z \in \mathbb{C} \text{ tel que } |z| > 1\}$ . Nous avons  $A \cup B = \mathbb{C} \setminus S^1$ . La partie  $A$  est connexe et même convexe. En ce qui concerne la connexité de  $B$ , c'est le corolaire 18.65.

De plus  $A$  et  $B$  sont des ouverts disjoints.

Supposons que  $C$  soit une composante connexe de  $\mathbb{C} \setminus S^1(0, 1)$  contenant un point de  $A$  et un point de  $B$ . Soit  $a \in C \cap A$ . Par définition 7.104(1) d'une composante connexe,  $C$  contient tous les connexes contenant  $a$ . En particulier  $A \subset C$ . De même si  $b \in C \cap B$ , alors  $B \subset C$ .

Dans ce cas nous aurions  $C = A \cup B = \mathbb{C} \setminus S^1(0, 1)$ . Mais cela est impossible parce que  $A$  et  $B$  forment un recouvrement de  $\mathbb{C} \setminus S^1$  par deux ouverts disjoints.

Donc  $A$  et  $B$  sont des composantes connexes de  $\mathbb{C} \setminus S^1$ . Et ce sont les deux seules parce que  $A \cup B = \mathbb{C} \setminus S^1$ .  $\square$

## 27.16.2 Théorème de Jordan

ThoHSPWBuh

**Théorème 27.216** (Théorème de Jordan[234, 768]).

Soit une courbe de Jordan  $\Gamma$  fermée dans  $\mathbb{C}$ .

ITEMooOAKXooMirWiD

(1) La partie  $\mathbb{C} \setminus \Gamma$  a exactement deux composantes connexes.

ITEMooICBGooXHMALK

(2) Une composante connexe est bornée, l'autre non.

(3) Les deux composantes connexes ont  $\Gamma$  comme frontière.

*Démonstration.* En plusieurs parties.

(1) **Pour (1)** Nous savons par le corolaire 27.196(2) que  $\Gamma$  est homéomorphe à  $S^1$ . Le corolaire 27.214 nous dit alors que  $\mathbb{C} \setminus \Gamma$  a le même nombre de composantes connexes que  $\mathbb{C} \setminus S^1$ , c'est-à-dire deux par le lemme 27.215.

(2) **Pour (2)** Le lemme 26.50 dit que  $\mathbb{C} \setminus \Gamma$  a exactement une composante connexe non bornée. Vu que nous venons de voir qu'elle a exactement deux composants connexes, c'est que l'autre est bornée.

 $\square$ 

## 27.17 Topologie faible

DEFooZGLDooRRarRj

**Définition 27.217** ([220, 1]).

Soit un espace vectoriel normé sur le corps  $\mathbb{K}$  ( $= \mathbb{R}$  ou  $\mathbb{C}$ )  $E$ , et son dual  $E'$ .

(1) Le dual  $E'$  a la topologie de la norme opérateur.

ITEMooBTIQooTQrtWt

(2) Le dual  $E'$  a également la topologie faible  $\sigma(E', E)$  donnée par la topologie initiale des fonctions d'évaluation

$$\begin{aligned} ev_x: E' &\rightarrow \mathbb{K} \\ \varphi &\mapsto \varphi(x). \end{aligned} \quad (27.1011)$$

EQooGGDFooSlyGGq

(3) L'espace  $E$  a lui-même une topologie faible  $\sigma(E, E')$  qui est la topologie initiale de tous les éléments de  $\varphi$ .

La **topologie faible** sur  $E$ , est la plus petite topologie  $\tau$  pour laquelle

$$(E, \|\cdot\|)' = (E, \tau)'. \quad (27.1012)$$

Autrement dit, c'est la topologie de la proposition 7.55 rendant continues toutes les applications de  $E'$ .

Elle sera notée  $\tau_w$  ou  $\sigma(E, E')$ .

LEMooZPJNooJyaeAC

**Lemme 27.218** ([1]).

Les ouverts faibles sont ouverts.

*Démonstration.* Nous avons  $\tau_w \subset \tau$  pour toute topologie  $\tau$  telle que  $(E, \tau_N)' = (E, \tau)'$  où  $\tau_N$  est la topologie de la norme. En particulier en posant  $\tau = \tau_N$  on a l'égalité. Donc  $\tau_w \subset \tau_N$ .  $\square$

Vu que, pour rendre continue une application, il suffit que toutes les images inverses des ouverts de  $\mathbb{F}$  soient des ouverts, la topologie faible sur  $E$  est également la topologie engendrée<sup>131</sup> par les parties  $\varphi^{-1}(V)$  avec  $\varphi \in E'$  et  $V \in \tau_{\mathbb{F}}$ .

**Remarque 27.219.**

Il faut noter que la topologie faible n'est pas une topologie métrique. Cela même si la condition  $A_i x \rightarrow Ax$ , elle, est métrique vu qu'elle est écrite dans  $E$ .

Dans le cas où  $E$  est de dimension infinie, la topologie faible est réellement différente de la topologie forte. Nous verrons à la sous-section 25.3.6 que dans le cas des projections sur un espace de Hilbert, l'égalité

$$\sum_{i=1}^{\infty} \text{proj}_{u_i} = \text{Id} \quad (27.1013)$$

est vraie pour la topologie faible, mais pas pour la topologie forte.

**Lemme 27.220** ([220]).

La topologie faible est Hausdorff<sup>132</sup>.

*Démonstration.* Soient un espace de Banach  $E$  ainsi que  $x_1 \neq x_2$  dans  $E$ . Les parties  $\{x_1\}$  et  $\{x_2\}$  vérifient les hypothèses de Hahn-Banach (seconde forme géométrique, 27.167). Il existe donc une fonctionnelle  $\varphi: E \rightarrow \mathbb{F}$  telle que  $\varphi(x_1) \neq \varphi(x_2)$ . Avec un peu de bonne volonté, nous supposons que  $\text{Re}(\varphi(x_1)) \neq \text{Re}(\varphi(x_2))$ . Soit  $\alpha \in \mathbb{R}$  tel que

$$\text{Re}(\varphi(x_1)) < \alpha < \text{Re}(\varphi(x_2)). \quad (27.1014)$$

Posons

$$\mathcal{O}_1 = \{x \in E \text{ tel que } \text{Re}(\varphi(x)) < \alpha\}, \quad (27.1015a)$$

$$\mathcal{O}_2 = \{x \in E \text{ tel que } \text{Re}(\varphi(x)) > \alpha\}. \quad (27.1015b)$$

Nous avons  $\mathcal{O}_i = \varphi^{-1}(V_i)$  avec<sup>133</sup>

$$V_1 = ]-\infty, \alpha[ + i\mathbb{R} \quad (27.1016)$$

et quelque chose du même genre pour  $V_2$ . Vu que  $V_1$  et  $V_2$  sont ouverts dans  $\mathbb{F}$ , et que  $\varphi$  est continue pour la topologie faible, les parties  $\mathcal{O}_i$  sont ouvertes dans  $(E, \tau_w)$ .

Nous avons de plus  $\mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$  et  $x_i \in \mathcal{O}_i$ , de telle sorte que  $x_1$  et  $x_2$  sont correctement séparés.  $\square$

131. Proposition 7.13.

132. Espace topologique Hausdorff, définition 7.74.

133. Adaptez si  $\mathbb{F} = \mathbb{R}$  au lieu de  $\mathbb{C}$ .

### 27.17.1 Bases de topologie

**Lemme 27.221** ([220]).

Soit un espace de Banach  $E$ . Soient  $x_0 \in E$ ,  $\varphi_1, \dots, \varphi_k \in E'$ . La partie

$$V(x_0; \varphi_1, \dots, \varphi_k, \epsilon) = \{x \in E \text{ tel que } |\varphi_i(x - x_0)| < \epsilon\} \quad (27.1017)$$

est faiblement ouverte.

*Démonstration.* Remarquez que

$$V(\varphi_1, \dots, \varphi_k, \epsilon) = \bigcap_{i=1}^k \varphi_i^{-1} \{z \in \mathbb{C} \text{ tel que } |z - \varphi_i(x_0)| < \epsilon\}. \quad (27.1018)$$

Vu que  $\varphi_i^{-1}$  transforme un ouvert en un ouvert,  $V(\varphi_1, \dots, \varphi_k, \epsilon)$  est une intersection d'ouverts, et donc un ouvert. De plus  $x_0$  est dedans parce que  $\varphi_i(x_0 - x_0) = \varphi_i(0) = 0 < \epsilon$ .  $\square$

**Lemme 27.222** ([220]).

Soient un espace de Banach  $E$  ainsi qu'un élément  $x_0 \in E$ . Les ensembles

$$V(x_0; \varphi_1, \dots, \varphi_k, \epsilon) = \{x \in E \text{ tel que } |\varphi_i(x - x_0)| < \epsilon \forall i\} \quad (27.1019)$$

forment une base de la topologie  $\sigma(E, E')$  en  $x_0$ .

*Démonstration.* Soit un ouvert faible  $U$  contenant  $x_0$ . Nous devons prouver que  $U$  contient une partie de la forme  $V(\varphi_1, \dots, \varphi_k, \epsilon)$ . C'est le moment d'avoir bien en tête la construction de la topologie engendrée donnée en la proposition 7.13.

La topologie faible est engendrée par les parties de la forme  $\varphi^{-1}(V)$  avec  $\varphi \in E'$  et  $V \in \tau_{\mathbb{F}}$ . La partie  $U$  est une union de parties de la forme

$$\bigcap_{j=1}^k \varphi_j^{-1}(V_j) \quad (27.1020)$$

où  $V_j$  est ouvert dans  $\mathbb{F}$  et  $\varphi_j \in E'$ . Vu que  $x_0 \in U$ , il existe une partie de la forme (27.1020) contenant  $x_0$ . Nous la notons  $W$  :

$$x_0 \in W = \bigcap_{j=1}^k \varphi_j^{-1}(V_j) \subset U. \quad (27.1021)$$

Vu que  $V_j$  est un ouvert de  $\mathbb{F}$  contenant  $\varphi_j(x_0)$ , il existe  $\delta_j > 0$  tel que  $B(\varphi_j(x_0), \delta_j) \subset V_j$ . En prenant  $\epsilon = \min_{j=1, \dots, k} (\delta_j)$ , nous avons  $B(\varphi_j(x_0), \epsilon) \subset V_j$  pour tout  $j$ . En particulier,

$$\bigcap_{j=1}^k \varphi_j^{-1} \left( B(\varphi_j(x_0), \epsilon) \right) \subset W. \quad (27.1022)$$

Nous montrons à présent que  $V(\varphi_1, \dots, \varphi_k, \epsilon) \subset W$ . Pour cela nous considérons  $x \in V(\varphi_1, \dots, \varphi_k, \epsilon)$  et nous montrons que  $x \in \varphi_j^{-1} \left( B(\varphi_j(x_0), \epsilon) \right)$  pour tout  $j = 1, \dots, k$ . Nous avons

$$|\varphi_j(x) - \varphi_j(x_0)| = |\varphi_j(x - x_0)| < \epsilon, \quad (27.1023)$$

et donc bien  $\varphi_j(x) \in B(\varphi_j(x_0), \epsilon)$ . Au final nous avons

$$V(\varphi_1, \dots, \varphi_k, \epsilon) \subset \bigcap_{j=1}^k \varphi_j^{-1} \left( B(\varphi_j(x_0), \epsilon) \right) \subset W \subset U. \quad (27.1024)$$

**Proposition 27.223** ([769]).

Soit un espace de Banach  $E$ . Soit  $\varphi_0 \in E'$ . Une base de la topologie<sup>134</sup>  $\sigma(E', E)$  en  $\varphi_0$  est donnée par les parties

$$V(\varphi_0; x_1, \dots, x_k, \epsilon) = \{\varphi \in E' \text{ tel que } |(\varphi - \varphi_0)(x_i)| < \epsilon \forall i = 1, \dots, k\}. \quad (27.1025)$$

134. Définition 27.217(2).

### 27.17.2 Convergence

**Proposition 27.224** ([220, 770]).

Soient un espace de Banach  $E$  ainsi qu'une suite  $(x_n)$  dans  $E$ .

- (1)  $x_n \xrightarrow{w} x$  si et seulement si  $\varphi(x_n) \rightarrow \varphi(x)$  pour tout  $\varphi \in E'$ .
- (2)  $x_n \xrightarrow{\|\cdot\|} x$  implique  $x_n \xrightarrow{w} x$ .
- (3) Si  $x_n \xrightarrow{w} x$  alors  $\{\|x_n\|\}$  est borné et

$$\|x\| \leq \liminf \|x_n\|. \quad (27.1026)$$

- (4) Si  $x_n \xrightarrow{w} x$  et  $\varphi_n \xrightarrow{E'} \varphi$  alors  $\varphi_n(x_n) \rightarrow \varphi(x)$ .

*Démonstration.* Point par point.

- (1) **Pour (1)** C'est le lemme 7.60.
- (2) **Pour (2)** Nous prouvons que  $\varphi(x_n) \rightarrow \varphi(x)$  pour tout  $\varphi \in E'$  :

$$|\varphi(x_n) - \varphi(x)| = |\varphi(x_n - x)| \leq \|\varphi\|_{E'} \|x_n - x\|_E \rightarrow 0 \quad (27.1027)$$

parce que par hypothèse  $\|x_n - x\| \rightarrow 0$ .

- (3) **Pour (3)**

Vu que  $\varphi(x_n) \xrightarrow{\mathbb{F}} \varphi(x)$ , l'ensemble  $\{\varphi(x_n)\}$  est borné dans  $\mathbb{F}$ . Considérons les opérateurs suivants :

$$\begin{aligned} T_n : E' &\rightarrow \mathbb{F} \\ \varphi &\mapsto \varphi(x_n). \end{aligned} \quad (27.1028)$$

Cet opérateur a la propriété que  $\|T_n\|_{(E')'} = \|x_n\|_E$  ; en effet, en utilisant la proposition 27.172,

$$\|T_n\| = \sup_{\|\varphi\|=1} |T_n(\varphi)| = \sup_{\|\varphi\|=1} |\varphi(x_n)| = \|x_n\|. \quad (27.1029)$$

Chacun des  $T_n$  est donc un opérateur linéaire borné. Nous vérifions que la famille  $\{T_n\}_{n \in \mathbb{N}}$  satisfait aux hypothèses du théorème de Banach-Steinhaus 10.174. D'abord  $E'$  est un Banach par la proposition 7.319. Ensuite nous venons de voir que chacun de  $T_n$  est borné dans  $E'$ . Et enfin, pour chaque  $\varphi \in E'$  nous avons

$$\sup_{n \in \mathbb{N}} \|T_n(\varphi)\| = \sup_{n \in \mathbb{N}} |\varphi(x_n)| < \infty. \quad (27.1030)$$

Le théorème de Banach-Steinhaus nous assure donc que

$$\sup_{n \in \mathbb{N}} \|x_n\| = \sup_{n \in \mathbb{N}} \|T_n\| < \infty. \quad (27.1031)$$

Donc l'ensemble  $\{\|x_n\|\}$  est borné.

Pour chaque  $\varphi \in E'$  tel que  $\|\varphi\| \leq 1$  nous avons

$$|\varphi(x_n)|_{\mathbb{F}} \leq \|\varphi\|_{E'} \|x_n\|_E \leq \|x_n\| \quad (27.1032)$$

Étant donnée l'inégalité  $|\varphi(x_n)| \leq \|x_n\|$  pour tout  $n$  et étant donnée la convergence  $|\varphi(x_n)| \rightarrow |\varphi(x)|$ , nous avons pour tout  $n$  :

$$|\varphi(x)| \leq \|x_n\|. \quad (27.1033)$$

Cette inégalité valable pour tout  $n$  donne la conclusion :

$$|\varphi(x)| \leq \liminf \|x_n\|. \quad (27.1034)$$

(4) **Pour (4)**

Nous avons  $\overline{\text{SUBEQooPHWLooZUiQLo}}$

$$|\varphi_n(x_n) - \varphi(x)| \leq |\varphi_n(x_n) - \varphi(x_n)| + |\varphi(x_n) - \varphi(x)| \quad (27.1035a)$$

$$\leq \|x_n\| \|\varphi_n - \varphi\| + |\varphi(x_n - x)| \quad (27.1035b)$$

Mais par hypothèse  $\|\varphi_n - \varphi\| \rightarrow 0$ , par le point (3),  $\|x_n\|$  est borné, et par le point (1), nous avons  $|\varphi(x_n) - \varphi(x)| \rightarrow 0$ . Tout ça mis ensemble nous permet de prendre la limite dans (27.1035) et de voir que  $\|\varphi_n(x_n) - \varphi(x)\| \rightarrow 0$ .

□

**Proposition 27.225.**

Si  $E$  est un espace de Banach de dimension finie, alors sa topologie normée est la même que sa topologie faible :  $\tau_w = \tau_{\|\cdot\|}$ .

*Démonstration.* Vu que  $E'$  est défini comme étant l'ensemble des formes continues pour la topologie  $\|\cdot\|$ , et que  $\tau_w$  est la plus petite topologie pour laquelle tous les éléments de  $E'$  sont continus, nous avons  $\tau_w \subset \tau_{\|\cdot\|}$ .

Pour prouver l'inclusion inverse, nous considérons  $U \in \tau_{\|\cdot\|}$  et nous prouvons que  $U$  est également un ouvert faible en montrant que tout élément de  $U$  est inclus dans un ouvert faible contenu dans  $U$ .

Soit  $x_0 \in U$  ainsi que  $r > 0$  tel que  $B(x_0, r) \subset U$ . Nous considérons une base  $\{e_i\}_{i=1, \dots, n}$  de  $E$  telle que  $\|e_i\| = 1$  pour tout  $i$ . Vu que tout élément de  $E$  peut être décomposé de façon unique en  $x = \sum_i x_i e_i$ , nous considérons les fonctionnelles linéaires

$$\begin{aligned} \varphi_j : E &\rightarrow \mathbb{F} \\ x &\mapsto x_j. \end{aligned} \quad (27.1036)$$

C'est le moment de ressortir notre ouvert préféré <sup>135</sup> autour de  $x_0$  :

$$V(x_0, \varphi_1, \dots, \varphi_n, \epsilon) = \bigcap_{i=1}^n \varphi_i^{-1} \left( \{ \|z\| < \epsilon \} \right). \quad (27.1037)$$

Supposons que  $x \in V(x_0, \varphi_1, \dots, \varphi_n, \epsilon)$ . Il vérifie  $|\varphi_j(x) - \varphi_j(x_0)| \leq \epsilon$  et donc

$$\|x - x_0\| = \left\| \sum_{j=1}^n \varphi_j(x - x_0) e_j \right\| \quad (27.1038a)$$

$$\leq \sum_j |\varphi_j(x - x_0)| \|e_j\| \quad (27.1038b)$$

$$= \sum_j |\varphi_j(x - x_0)| \quad (27.1038c)$$

$$\leq \sum_j |\varphi_j(x) - \varphi_j(x_0)| \quad (27.1038d)$$

$$\leq n\epsilon. \quad (27.1038e)$$

En choisissant  $\epsilon < \frac{r}{n}$ , nous avons  $x \in B(x_0, r)$  et donc

$$V(x_0, \dots) \subset B(x_0, r). \quad (27.1039)$$

□

---

135. Nous avons un ...faible pour lui !



LEMooMCYAooGMzbbs

**Lemme 27.226.**

Soit un espace de Banach  $E$  sur le corps  $\mathbb{F}$ . Nous notons  $\text{Fun}(E, \mathbb{F})$  l'ensemble de toutes les applications de  $E$  vers  $\mathbb{F}$ . Pour chaque  $x \in E$  nous considérons l'application

$$\begin{aligned} f_x: \text{Fun}(E, \mathbb{F}) &\rightarrow \mathbb{F} \\ \omega &\mapsto \omega(x). \end{aligned} \quad (27.1040)$$

Nous considérons sur  $\text{Fun}(E, \mathbb{F})$  la plus petite topologie telle que tous les  $f_x$  soient continues.

La suite  $(\omega_n)$  dans  $\text{Fun}(E, \mathbb{F})$  converge vers  $\omega \in \text{Fun}(E, \mathbb{F})$  si et seulement si  $\omega_n(x) \rightarrow \omega(x)$  pour tout  $x \in E$ .

LEMooWXBVooSjafZr

**Lemme 27.227.**

Soit un espace de Banach  $E$  sur le corps  $\mathbb{F}$  ( $= \mathbb{R}$  ou  $\mathbb{C}$ ). Nous considérons l'ensemble  $\text{Fun}(E, \mathbb{F})$  sur lequel nous mettons la topologie minimale qui rend continue les applications

$$\begin{aligned} f_x: \text{Fun}(E, \mathbb{F}) &\rightarrow \mathbb{F} \\ \omega &\mapsto \omega(x). \end{aligned} \quad (27.1041)$$

Nous notons  $I = \{f_x \text{ tel que } x \in E\}$  et  $\tau_I$  la topologie de  $\text{Fun}(E, \mathbb{F})$ .

Si  $(\omega_n)$  est une suite dans  $\text{Fun}(E, \mathbb{F})$  nous avons  $\omega_n \xrightarrow{\tau_I} \omega$  si et seulement si  $\omega_n(x) \rightarrow \omega(x)$  pour tout  $x \in E$ .

THOooRECTooEVLHSq

**Théorème 27.228** (Banach-Alaoglu-Bourbaki[220]).

Soit un espace de Banach  $E$ . La boule unité

$$B = \{\varphi \in E' \text{ tel que } \|\varphi\| = 1\} \quad (27.1042)$$

est compacte pour la topologie faible.

*Démonstration.* Nous considérons à nouveau  $\text{Fun}(E, \mathbb{F})$  muni de la topologie du lemme 27.227. Nous considérons l'inclusion

$$\begin{aligned} \Phi: (E', \|\cdot\|) &\rightarrow (\text{Fun}(E, \mathbb{F}), \tau_I) \\ \omega &\mapsto \omega. \end{aligned} \quad (27.1043)$$

En particulier nous notons  $S = \Phi(E')$  et nous allons prouver que  $\Phi: E' \rightarrow S$  bijective et continue et que  $\Phi^{-1}: S \rightarrow E'$  est également continue.  $\square$

PROPooYARHooOpmztY

**Proposition 27.229** ([220]).

Un convexe dans un espace de Banach est fermé si et seulement si il est faiblement<sup>136</sup> fermé.

*Démonstration.* En deux parties.

- (1)  $\Rightarrow$  Soit un connexe fermé  $C$  de  $E$ . Nous devons prouver que  $C$  est faiblement fermé, c'est-à-dire que  $E \setminus C$  est faiblement ouvert. Soit  $x_0 \in E \setminus C$ .

Étant donné que tout espace de Banach est localement convexe (proposition 7.255) Nous utilisons le théorème de Hahn-Banach 27.167 pour les parties  $A = \{x_0\}$  (qui est compacte) et  $C = C$  (qui est fermé). Il existe donc un hyperplan qui sépare strictement<sup>137</sup>  $\{x_0\}$  et  $C$ . Autrement dit, il existe  $\alpha \in \mathbb{R}$  et  $f \in E'$  tels que

$$f(x_0) < \alpha < f(x) \quad (27.1044)$$

pour tout  $x \in C$ .

Nous posons  $V = \{x \in E \text{ tel que } f(x) < \alpha\}$ , qui n'est autre que le  $V(f, \alpha)$  du lemme 27.221. C'est un ouvert faible contenant  $x_0$  et contenu dans  $E \setminus C$ . Cela prouve que  $E \setminus C$  est faiblement ouvert.

136. Topologie faible, définition 27.217.

137. Hyperplan qui sépare, définition 27.164.

- (2)  $\Leftarrow$  Nous supposons que  $C$  est faiblement fermé. Donc  $E \setminus C$  est faiblement ouvert et donc ouvert (lemme 27.218). Donc  $C$  est fermé.

□

PROPooGQLOooNSgZnZ

**Proposition 27.230** ([220, 237]).

Soit un espace de Banach  $E$  sur  $\mathbb{K}$  ( $\mathbb{R}$  ou  $\mathbb{C}$ ) et un sous-espace vectoriel fermé  $M$  de  $E$ .

- (1)  $M$  muni de la norme hérité de  $E$  est un espace de Banach.  
 (2) Nous considérons les deux topologies suivantes sur  $M$  :  
 —  $\tau_1$  : la topologie induite depuis  $\sigma(E, E')$ .  
 —  $\tau_2$  : la topologie  $\sigma(M, M')$ .

Alors  $\tau_1 = \tau_2$ .

*Démonstration.* En deux parties.

- (1) **Espace de Banach** C'est la proposition 7.318.

- (2) **Égalité des topologies** En deux parties.

- (2a)  $\tau_2 \subset \tau_1$  Nous montrons que  $\tau_1$  rend continue tous les éléments de  $M'$ . Soit  $\varphi \in M'$ . Soit un ouvert  $U$  de  $\mathbb{K}$  ; nous devons prouver que  $\varphi^{-1}(U) \in \tau_1$ . Pour cela, le théorème de Hahn-Banach 27.171 nous indique qu'il existe  $\tilde{\varphi} \in E'$  tel que  $\varphi = \tilde{\varphi}|_M$ . Vu que  $\sigma(E, E')$  rend continue tous les éléments de  $E'$ , nous avons

$$\tilde{\varphi}^{-1}(U) \in \sigma(E, E'). \quad (27.1045)$$

Donc

$$\varphi^{-1}(U) = \tilde{\varphi}^{-1}(U) \cap M \in \tau_1. \quad (27.1046)$$

Nous en déduisons que  $\tau_1$  rend continus tous les éléments de  $M'$ . Vu que  $\sigma(M, M')$  est la plus petite faisant ça, nous avons bien  $\tau_2 \subset \tau_1$ .

- (2b)  $\tau_1 \subset \tau_2$  Soit  $\mathcal{O} \in \tau_1$ . Nous avons  $\mathcal{O} = U \cap M$  avec  $U \in \sigma(E, E')$ . Soit  $x_0 \in \mathcal{O}$ . Étant donné que  $x_0 \in U$  et étant donné le lemme 27.222, il existe  $\varphi_1, \dots, \varphi_k \in E'$  et  $\epsilon > 0$  tels que

$$V(x_0; \varphi_1, \dots, \varphi_k, \epsilon) = \{x \in E \text{ tel que } |\varphi_i(x - x_0)| < \epsilon\} \subset U. \quad (27.1047)$$

En notant  $f_i$  la restriction de  $\varphi_i$  à  $M$ , nous avons

$$\{x \in E \text{ tel que } |\varphi_i(x - x_0)| < \epsilon\} \cap M = \{x \in M \text{ tel que } |\varphi_i(x - x_0)| < \epsilon\} \quad (27.1048a)$$

$$= \{x \in M \text{ tel que } |f_i(x - x_0)| < \epsilon\} \quad (27.1048b)$$

$$\in \sigma(M, M'). \quad (27.1048c)$$

Cette partie est donc un  $\sigma(M, M')$ -voisinage de  $x_0$  contenu dans  $\mathcal{O}$ . Donc  $\mathcal{O} \subset \tau_2$  et donc  $\tau_1 \subset \tau_2$ .

□

### 27.17.3 Helly et Goldstine

LEMooSBWJooKIfuJj

**Lemme 27.231** (Helly[220, 771]).

Soit un espace de Banach  $E$  sur  $\mathbb{K}$  ( $= \mathbb{R}$  ou  $\mathbb{C}$ ). Nous considérons  $f_1, \dots, f_k \in E'$  ainsi que des éléments  $(\gamma_1, \dots, \gamma_k) \in \mathbb{K}^n$ . De plus nous introduisons

$$\begin{aligned} \varphi: E &\rightarrow \mathbb{K}^k \\ x &\mapsto (f_1(x), \dots, f_k(x)). \end{aligned} \quad (27.1049)$$

Nous avons équivalence entre

ITEMooDMOOooPJtgTE

- (1) Pour tout  $\epsilon > 0$ , il existe  $x_\epsilon \in E$  tel que  $\|x_\epsilon\| < 1$  et  $|f_i(x_\epsilon) - \gamma_i| < \epsilon$ .

ITEMooAQAJooHmSXGW

$$(2) \quad \gamma \in \overline{\varphi(B(0, 1))}.$$

ITEMooXJKAAoFHuSOv

$$(3) \quad |\sum_{l=1}^k \beta_l \gamma_l| \leq \|\sum_{l=1}^k \beta_l f_l\| \text{ pour tout } \beta_1, \dots, \beta_k \in \mathbb{K}.$$

Notez que  $|\cdot|$  dénote soit la valeur absolue (si  $\mathbb{K} = \mathbb{R}$ ) soit le module d'un nombre complexe (si  $\mathbb{K} = \mathbb{C}$ ). Par ailleurs  $\|\cdot\|$  dénote la norme opérateur.

*Démonstration.* En plusieurs parties.

$$(1) \quad \underline{(1) \Rightarrow (2)} \quad \text{Soit } \epsilon > 0. \text{ Il existe } x_\epsilon \in E \text{ tel que } x_\epsilon \in \overline{B(0, 1)}, \text{ et donc tel que}$$

$$\|\varphi(x_\epsilon) - \gamma\| < k\epsilon \quad (27.1050)$$

où à gauche, c'est la norme dans  $\mathbb{K}^k$ . Cela prouve que, en utilisant 7.50,

$$\gamma \in \varphi(\overline{B(0, 1)}) \subset \overline{\varphi(B(0, 1))} \quad (27.1051)$$

$$(2) \quad \underline{(2) \Rightarrow (1)} \quad \text{Soit } \epsilon > 0. \text{ Étant donné que } \gamma \in \overline{\varphi(B(0, 1))}, \text{ il existe } y_\epsilon \in \varphi(B(0, 1)) \text{ tel que } \|y_\epsilon - \gamma\| < \epsilon \text{ et donc } x_\epsilon \in B(0, 1) \text{ tel que } \|\varphi(x_\epsilon) - \gamma\| < \epsilon.$$

Par inégalité triangulaire, cela implique  $|f_l(x_\epsilon) - \gamma_l| < \epsilon$ .

$$(3) \quad \underline{(1) \Rightarrow (3)} \quad \text{Nous posons } S = \sum_{i=1}^k |\beta_i|. \text{ Soit } \epsilon > 0. \text{ Nous avons}$$

$$|\sum_i \beta_i f_i(x_\epsilon) - \sum_i \beta_i \gamma_i| \leq |\sum_i \beta_i (f_i(x_\epsilon) - \gamma_i)| \quad (27.1052a)$$

$$\leq \sum_i |\beta_i| |f_i(x_\epsilon) - \gamma_i| \quad (27.1052b)$$

$$\leq S\epsilon. \quad (27.1052c)$$

Grâce à la formule de la proposition 10.2, nous en déduisons

$$|\sum_i \beta_i \gamma_i| \leq |\sum_i \beta_i f_i(x_\epsilon)| + S\epsilon \quad (27.1053a)$$

$$\leq \|\sum_i \beta_i f_i\| \|x_\epsilon\| + S\epsilon \quad (27.1053b)$$

$$\leq \|\sum_i \beta_i f_i\| + S\epsilon. \quad (27.1053c)$$

Vu que cette inégalité est valide pour tout  $\epsilon$ , nous déduisons que  $|\sum_i \beta_i \gamma_i| \leq \|\sum_i \beta_i f_i\|$ .

$$(4) \quad \underline{(3) \Rightarrow (1)} \quad \text{Nous faisons non(1) } \Rightarrow \text{ non(3). Et comme (1) est équivalent à (2), nous faisons non(2) } \Rightarrow \text{ non(3).}$$

Nous avons donc  $\gamma \notin \overline{\varphi(B(0, 1))}$ . En utilisant le théorème 27.167 ou 27.167 selon que  $\mathbb{K} = \mathbb{C}$  ou  $\mathbb{K} = \mathbb{R}$ , nous savons que  $\{\gamma\}$  et  $\overline{\varphi(B(0, 1))}$  sont strictement séparés par un hyperplan de  $\mathbb{K}^k$ . Autrement dit, il existe  $\beta \in \mathbb{K}^k$  et  $\alpha \in \mathbb{R}$  tels que

$$\operatorname{Re}(\beta \cdot \varphi(x)) < \alpha < \operatorname{Re}(\beta \cdot \gamma). \quad (27.1054)$$

Autrement dit,

$$\operatorname{Re}(\sum_{i=1}^k \beta_i f_i(x)) < \alpha < \operatorname{Re}(\sum_i \beta_i \gamma_i). \quad (27.1055)$$

EQooSEURooAoxMVa

Maintenant on fait deux cas selon que  $\mathbb{K} = \mathbb{C}$  ou  $\mathbb{K} = \mathbb{R}$ .

$$(4a) \quad \underline{\text{Si } \mathbb{K} = \mathbb{R}} \quad \text{Alors nous pouvons enlever le Re dans (27.1055). Vu que l'inégalité est valide pour tout } x \in E \text{ et vu que } |A(-x)| = |A(x)| \text{ lorsque } A \text{ est linéaire, nous avons}$$

$$|(\sum_i \beta_i f_i)(x)| < \alpha < \sum_i \beta_i \gamma_i. \quad (27.1056)$$

En prenant le suprénum sur  $\|x\| = 1$  nous trouvons

$$\left\| \sum_i \beta_i f_i \right\| \leq \alpha \leq \sum_i \beta_i \gamma_i, \quad (27.1057)$$

qui contredit (3).

- (4b) **Si  $\mathbb{K} = \mathbb{C}$**  Nous ne pouvons pas enlever les modules si facilement. C'est d'ailleurs le moment de relire le lemme 27.159. Nous posons  $\phi = \sum_i \beta_i f_i$  et  $g = \operatorname{Re}(\phi)$ . Ce que dit (27.1055) est

$$|g(x)| < \alpha < \operatorname{Re}\left(\sum_i \beta_i \gamma_i\right) \leq \left|\sum_i \beta_i \gamma_i\right|. \quad (27.1058)$$

En prenant le suprénum sur  $\|x\| = 1$  et en tenant compte du fait que  $\|\phi\| = \|g\|$ , nous obtenons à nouveau une inégalité similaire à (27.1057), et donc une contradiction.

□

**Théorème 27.232** (Théorème de Goldstine[771]).

Soit un espace de Banach  $E$  sur  $\mathbb{K}$  ( $= \mathbb{R}$  ou  $\mathbb{C}$ ). Nous avons<sup>138</sup>

- (1) La partie  $J(B_E(0, 1))$  est dense dans  $B_{E''}(0, 1)$  pour la topologie faible<sup>139</sup>  $\sigma(E'', E')$ .  
 (2) La partie  $J(E)$  est  $\sigma(E'', E')$ -dense dans  $E''$ .

*Démonstration.* En deux parties.

- (1) **Pour (1)** Soit  $\xi_0 \in B_{E''}(0, 1)$ , et soit un voisinage  $V$  de  $\xi_0$ . Nous devons prouver que  $V \cap J(B_E(0, 1)) \neq \emptyset$ . La proposition 27.223 donne une base de la topologie faible. Nous pouvons donc supposer que  $V$  est de la forme (27.1025). Il faut juste ajouter des primes un peu partout parce que nous parlons ici de  $\sigma(E', E'')$  et non de  $\sigma(E, E')$ . Bref. Il existe des  $\varphi_i \in E'$  et  $\epsilon > 0$  tels que

$$V = V(\xi_0; \varphi_1, \dots, \varphi_k, \epsilon) = \{\eta \in E'' \text{ tel que } |(\eta - \xi_0)(\varphi_i)| < \epsilon\}. \quad (27.1059)$$

Nous allons montrer qu'il existe  $x_\epsilon \in B_E(0, 1)$  tel que  $J(x_\epsilon) \in V$ .

Nous posons  $\gamma_i = \xi_0(\varphi_i) \in \mathbb{K}$ . Soient  $\beta_1, \dots, \beta_k \in \mathbb{K}$ . Nous avons

$$\left| \sum_{i=1}^k \beta_i \gamma_i \right| = \left| \sum_i \beta_i \xi_0(\varphi_i) \right| = |\xi_0(\sum_i \beta_i \varphi_i)| \leq \left\| \sum_i \beta_i \varphi_i \right\| \quad (27.1060)$$

parce que  $\|\xi_0\| \leq 1$ . L'inégalité (27.1060) étant valide pour tout choix de  $\beta_i$ , le théorème d'Helly 27.231 implique l'existence de  $x_\epsilon \in E$  tel que  $\|x_\epsilon\| < 1$  et

$$|\varphi_i(x_\epsilon) - \gamma_i| < \epsilon. \quad (27.1061)$$

Nous avons donc

$$\epsilon < |\varphi_i(x_\epsilon) - \gamma_i| = |\varphi_i(x_\epsilon) - \xi_0(\varphi_i)| = |(J(x_\epsilon) - \xi_0)(\varphi_i)|, \quad (27.1062)$$

et donc  $J(x_\epsilon) \in V$ .

- (2) **Pour (2)** Soit  $\xi_0 \in E''$  ainsi qu'un voisinage  $V$  de  $\xi_0$  de la forme

$$V(\xi_0; \varphi_1, \dots, \varphi_k, \epsilon) = \{\eta \in E'' \text{ tel que } |(\eta - \xi_0)(\varphi_i)| < \epsilon\}. \quad (27.1063)$$

Il existe  $\lambda \in \mathbb{K}$  tel que  $\lambda \xi_0 \in B_{E''}(0, 1)$ . Nous avons

$$\lambda V(\xi_0; \varphi_1, \dots, \varphi_k, \epsilon) = \{\lambda \eta \text{ tel que } |(\eta - \xi_0)(\varphi_i)| < \epsilon\} \quad (27.1064a)$$

$$= \{\lambda \eta \text{ tel que } |(\lambda \eta - \lambda \xi_0)(\varphi_i)| < \lambda \epsilon\} \quad (27.1064b)$$

$$= \{\eta \text{ tel que } |(\eta - \lambda \xi_0)(\varphi_i)| < \lambda \epsilon\}, \quad (27.1064c)$$

138. Définition 27.179 pour  $J$ .

139. Définie en 27.217.

autrement dit

$$\lambda V(\xi_0; \varphi_1, \dots, \varphi_k, \epsilon) = V(\lambda \xi_0; \varphi_1, \dots, \varphi_k, \lambda \epsilon). \quad (27.1065)$$

Étant donné que  $\lambda \xi_0 \in B_{E''}(0, 1)$ , la première partie dit qu'il existe  $x_\epsilon \in B_E(0, 1)$  tel que

$$J(x_\epsilon) \in V(\lambda \xi_0; \varphi_1, \dots, \varphi_k, \lambda \epsilon) = \lambda V(\xi_0; \varphi_1, \dots, \varphi_k, \epsilon). \quad (27.1066)$$

En reprenant (27.1065) (avec  $1/\lambda$  au lieu de  $\lambda$ ), nous avons

$$J(x_\epsilon/\lambda) \in \frac{1}{\lambda} V(\lambda \xi_0; \varphi_1, \dots, \varphi_k, \epsilon) = V(\xi_0; \varphi_1, \dots, \varphi_k, \epsilon), \quad (27.1067)$$

ce qui signifie que  $V(\xi_0; \varphi_1, \dots, \varphi_k, \epsilon)$  contient bien un élément de  $J(E)$ . □

## 27.17.4 Banach-Alaoglu

**Lemme 27.233** ([1, 772]).

Soit un espace vectoriel normé  $E$  sur  $\mathbb{K}$  ( $\mathbb{R}$  ou  $\mathbb{C}$ ). Pour chaque  $x \in E$  nous considérons  $K_x = \overline{B_{\mathbb{K}}(0, \|x\|)}$ . Nous considérons ensuite <sup>140</sup>  $K = \prod_{x \in E} K_x$ .

L'ensemble  $K$  a la topologie produit <sup>141</sup>  $\tau_K$  et le dual  $E'$  a la topologie faible  $\sigma(E', E)$ .

- (1) Nous avons  $K \cap E' = \overline{B_{E'}(0, 1)}$ .
- (2) Si nous désignons par  $\tau_P$  la topologie sur  $\overline{B_{E'}(0, 1)}$  induite depuis  $K$  et par  $\tau_\sigma$  celle induite depuis  $E'$ , alors nous avons  $\tau_P = \tau_\sigma$ .

*Démonstration.* En plusieurs parties.

- (1)  $K \cap E' = \overline{B_{E'}(0, 1)}$  Dans un premier sens, si  $f \in K \cap E'$ , alors

$$\frac{|f(x)|}{\|x\|} \leq \frac{\|x\|}{\|x\|} = 1, \quad (27.1068)$$

et donc  $f \in \overline{B_{E'}(0, 1)}$ . Dans l'autre sens, si  $f \in \overline{B_{E'}(0, 1)}$ , alors  $\|f(x)\| \leq \|x\|$  pour tout  $x$ , de telle sorte que  $f(x) \in K_x$  pour tout  $x$ , c'est-à-dire  $f \in K$ .

- (2) Base de topologie de  $\tau_K$  sur  $K$  En tant que topologie produit, la topologie  $\tau_K$  est la topologie initiale des projections

$$\begin{aligned} \pi_x: \prod_{y \in E} K_y &\rightarrow K_x \\ f &\mapsto f(x). \end{aligned} \quad (27.1069)$$

La proposition 7.33 donne une base de topologie sur  $K_x$  sous la forme de parties de  $\mathbb{K}$  de la forme

$$A_x(z, r) = B_{\mathbb{K}}(z, r) \cap B_{\mathbb{K}}(0, \|x\|). \quad (27.1070)$$

Ensuite le lemme 7.56 dit qu'une base de topologie de  $K$  est formée de parties de la forme

$$\bigcap_{k=1}^n \pi_{x_i}^{-1}(A_{x_i}(z_i, r_i)). \quad (27.1071)$$

- (3)  $\tau_\sigma \subset \tau_P$  Nous considérons un élément de  $\tau_\sigma$  sous la forme  $U' = \overline{B_{E'}(0, 1)} \cap V$  où  $V$  est un élément de  $\sigma(E', E)$ . Notre objectif est de montrer que c'est un élément de  $\tau_P$ . Pour cela nous allons prendre un élément  $\varphi_0 \in U'$  et montrer qu'il existe un voisinage  $W \in \tau_P$  tel que  $\varphi_0 \in W \subset U'$ .

140. Définition 1.22 pour le produit cartésien infini.

141. Topologie produit infinie, définition 7.62.

Par la proposition 27.223 nous considérons des points  $x_1, \dots, x_k$  et  $\epsilon > 0$  tels que  $V(\varphi_0, x_1, \dots, x_k, \epsilon) \subset V$ , et nous notons

$$U = \overline{B_{E'}(0, 1)} \cap V(\varphi_0, x_1, \dots, x_k, \epsilon). \quad (27.1072)$$

Nous construisons  $W$  de la façon suivante. Nous posons

$$A_{x_i} = B_{\mathbb{K}}(\varphi_0(x_i), \epsilon), \quad A_x = B_{\mathbb{K}}(0, \|x\|), \quad (27.1073)$$

la seconde définition étant celle lorsque  $x$  n'est pas un des  $x_1, \dots, x_k$ . Nous définissons ensuite

$$W = \overline{B_{E'}(0, 1)} \cap \prod_{x \in E} A_x. \quad (27.1074)$$

Avec ça nous avons bien  $\varphi_0 \in W$  parce que  $|\varphi_0(x)| \leq \|x\|$  parce que  $\varphi_0 \in E'$ . De plus  $W \subset U$ . En effet si  $h \in W$  nous avons

$$h(x_i) \in B_{\mathbb{K}}(\varphi_0(x_i), \epsilon), \quad (27.1075)$$

ce qui signifie que  $|h(x_i) - \varphi_0(x_i)| < \epsilon$ , et donc que  $h \in U$ .

- (4)  $\tau_P \subset \tau_\sigma$  Nous considérons un élément de  $\tau_P$  de la forme  $U = \overline{B_{E'}(0, 1)} \cap U_0$  où  $U_0$  est un élément de la topologie produit<sup>142</sup> de  $\prod_{x \in E} K_x$ . Soit  $\varphi_0 \in U$ .

Nous allons montrer qu'il existe  $W \in \tau_\sigma$  tel que  $\varphi_0 \in W \subset U$ . Nous cherchons  $W$  sous la forme  $W = \overline{B_{E'}(0, 1)} \cap W_0$  avec  $W_0 \in \sigma(E', E)$ .

Nous avons vu dans le point (2) qu'il existe  $z_i \in \mathbb{K}$ ,  $r_i > 0$  et  $x_i \in E$  tels que, en posant

$$A_i(z_i, r_i) = B_{\mathbb{K}}(z_i, r_i) \cap B_{\mathbb{K}}(0, \|x_i\|), \quad (27.1076)$$

nous avons

$$\varphi_0 \in U'_0 = \bigcap_{i=1}^k \pi_{x_i}^{-1}(A_i(z_i, r_i)) \subset U_0. \quad (27.1077)$$

Notez que

$$\pi_{x_i}^{-1}(A_i(z_i, r_i)) = \{f: E \rightarrow \mathbb{K} \text{ tel que } |f(x_i) - z_i| < r_i \text{ et } |f(x_i)| < \|x_i\|\}. \quad (27.1078)$$

Vu que  $\varphi_0 \in U'_0$ , il existe  $\alpha_i > 0$  tel que  $|\varphi_0(x_i) - z_i| < \alpha_i < r_i$ . Donc en prenant  $\epsilon_i < r_i - \alpha_i$ , nous avons

$$\{f: E \rightarrow \mathbb{K} \text{ tel que } |f(x_i) - \varphi_0(x_i)| < \epsilon_i\} \subset \{f: E \rightarrow \mathbb{K} \text{ tel que } |f(x_i) - z_i| < r_i\}. \quad (27.1079)$$

Tout ça pour dire qu'en posant

$$D_i = \{f: E \rightarrow \mathbb{K} \text{ tel que } |f(x_i) - \varphi_0(x_i)| < \epsilon_i \text{ et } |f(x_i)| \leq \|x_i\|\}, \quad (27.1080)$$

et  $U''_0 = \bigcap_{i=1}^k D_i$ , nous avons que  $U''_0 \in \tau_K$  et

$$\varphi_0 \in U''_0 \subset U'_0 \subset U_0. \quad (27.1081)$$

En prenant les mêmes  $\epsilon_i$ , nous posons

$$W_i = \{\eta \in E' \text{ tel que } |\eta(x_i) - \varphi_0(x_i)| < \epsilon_i\}, \quad (27.1082)$$

et  $W_0 = \bigcup_{i=0}^k W_i$ . Nous avons

$$\varphi_0 \in W_0 \cap \overline{B_{E'}(0, 1)} \subset U''_0 \cap \overline{B_{E'}(0, 1)} \subset U_0 \cap \overline{B_{E'}(0, 1)}. \quad (27.1083)$$

□

142. Définition 7.62.

THOooJNlQooOXGCKY

**Théorème 27.234** (Banach-Alaoglu[773]).

Soit un espace vectoriel normé  $E$  sur  $\mathbb{K}$  ( $\mathbb{R}$  ou  $\mathbb{C}$ ). Alors la boule fermée <sup>143</sup>  $\overline{B_{E'}(0,1)}$  est compacte pour la topologie faible  $\sigma(E', E)$ .

*Démonstration.* Pour chaque  $x \in E$ , nous posons  $K_x = [-\|x\|, \|x\|]$  et nous considérons le produit cartésien <sup>144</sup>  $K = \prod_{x \in E} K_x$ . Un élément de  $K$  est une application  $f: E \rightarrow \mathbb{K}$  vérifiant  $|f(x)| \leq \|x\|$ .

- (1) **Une topologie sur  $\text{Fun}(E, \mathbb{K})$**  Sur l'ensemble  $\text{Fun}(E, \mathbb{K})$  de toutes les applications  $E \rightarrow \mathbb{K}$  nous considérons la topologie initiale des applications

$$\begin{aligned} \pi_x: \text{Fun}(E, \mathbb{K}) &\rightarrow \mathbb{K} \\ f &\mapsto f(x). \end{aligned} \quad (27.1084)$$

C'est une topologie séparée <sup>145</sup> par la proposition 7.58. Nous la nommons  $\tau_F$

- (2)  **$K$  est compact dans  $\text{Fun}(E, \mathbb{K})$**  Pour chaque  $x \in E$ , la partie  $K_x$  est compacte. Donc la proposition 7.361 dit que  $K = \prod_{x \in E} K_x$  est compacte dans  $\text{Fun}(E, \mathbb{K}) = \prod_{x \in E} \mathbb{K}$ .
- (3) **Topologies sur  $\overline{B_{E'}(0,1)}$**  Nous avons vu dans le lemme 27.233 que  $\overline{B_{E'}(0,1)} = K \cap E'$ . Nous pouvons donc considérer (au moins) deux topologies dessus : celle induite de  $K$  et celle induite de  $E'$ . Nous nommons  $\tau_P$  la topologie de  $\overline{B_{E'}(0,1)}$  induite de  $K$  et  $\tau_\sigma$  celle induite de  $E'$ .

Notons que  $\tau_P$  est induite de  $\tau_K$  qui est induite de  $\tau_F$ . Donc  $\tau_P$  est la topologie induite de  $\text{Fun}(E, \mathbb{K})$  sur  $\overline{B_{E'}(0,1)}$ .

Toujours d'après le lemme 27.233, nous avons  $\tau_P = \tau_\sigma$ .

- (4)  **$E'$  est fermé dans  $K$**  Nous montrons que  $K \setminus E'$  est ouvert. Soit  $f \in K \setminus E'$ . Il y a deux possibilités : soit il existe  $x, y \in E$  tels que  $f(x) + f(y) \neq f(x+y)$ , soit il existe  $x \in E$  et  $\lambda \in \mathbb{K}$  tels que  $f(\lambda x) \neq \lambda f(x)$ .

- (4a) **Si  $f(x) + f(y) \neq f(x+y)$**  Il existe  $r > 0$  tel que

$$|f(x) + f(y) - f(x+y)| > r. \quad (27.1085)$$

Nous considérons cet ouvert de  $K$  :

$$A = \{g: E \rightarrow \mathbb{K} \text{ tel que } \begin{cases} g \in K, \\ |f(x) - g(x)| < \epsilon, \\ |f(y) - g(y)| < \epsilon, \\ |f(x+y) - g(x+y)| < \epsilon \end{cases} \}. \quad (27.1086)$$

Si  $g \in A$  nous avons :

$$\left| (g(x) + g(y) - g(x+y)) - (f(x) + f(y) - f(x+y)) \right| \quad (27.1087a)$$

$$\leq |g(x) - f(x)| + |g(y) - f(y)| \quad (27.1087b)$$

$$+ |g(x+y) - f(x+y)|$$

$$\leq 3\epsilon. \quad (27.1087c)$$

En choisissant  $\epsilon < r/3$  nous avons alors  $|g(x) + g(y) - g(x+y)| > 0$ , et donc  $g \in K \setminus E'$ .

143. Fermée par rapport à la topologie de la norme opérateur.

144. Définition 1.22.

145. Je n'utilise pas le fait que ce soit séparé dans la suite alors que [773] insiste dessus. À mon avis c'est parce que le corolaire 7.122 est souvent énoncé pour des espaces séparés. Peut-être parce que les Français disent que les compacts doivent être séparés. Dans le Frido, un compact doit seulement vérifier la propriété de sous-recouvrement fini. On peut être compact sans être séparé. Si vous avez une idée sur l'endroit où la séparation de  $\text{Fun}(E, \mathbb{K})$  joue dans la preuve, écrivez-moi.

- (4b) **Si**  $f(\lambda x) \neq \lambda f(x)$  Le même type de raisonnement <sup>146</sup> donne un ouvert de  $K$  contenant  $f$  et contenu dans  $K \setminus E'$ .
- (4c) **Conclusion**  $E'$  **est fermé** Contenant un voisinage de tout élément, la partie  $K \setminus E'$  est ouverte <sup>147</sup>. Donc  $E'$  est fermé.
- (5) **Conclusions** Dans  $\text{Fun}(E, \mathbb{K})$ , la partie  $K$  est compacte et la partie  $E'$  est fermée. Donc  $K \cap E'$  est compact pour la topologie  $\tau_F$  (corolaire 7.122). Donc  $(\overline{B_{E'}(0, 1)}, \tau_P)$  est un espace topologique compact.
- Nous avons vu que  $\tau_P = \tau_\sigma$ . Donc  $(\overline{B_{E'}(0, 1)}, \tau_\sigma)$  est un espace topologique compact. La proposition 7.31 dit alors que  $\overline{B_{E'}(0, 1)}$  est compact dans  $(E', \sigma(E', E))$ .

□

### 27.17.5 Espace de Banach réflexif, Kakutani

THOooTFIHooPQjVAr

**Théorème 27.235** (Kakutani[220]).

Un espace de Banach est réflexif<sup>148</sup> si et seulement si la boule fermée  $\overline{B_E(0, 1)}$  est compacte dans  $(E, \sigma(E, E'))$ <sup>149</sup>.

*Démonstration.* Deux sens.

- (1)  $\Rightarrow$  Nous supposons que  $E$  est un espace de Banach réflexif. Nous commençons par voir que

$$J^{-1}: (E'', \sigma(E'', E'))' \rightarrow (E, \sigma(E, E')) \quad (27.1088)$$

est continue. Nous utilisons la proposition 7.61 : l'application  $J^{-1}$  sera continue si et seulement si  $\varphi \circ J^{-1}: E'' \rightarrow \mathbb{K}$  est continue pour tout  $\varphi \in E'$ . Pour s'en assurer nous commençons par trouver une forme pratique pour  $\varphi \circ J^{-1}$ . Soit  $\xi \in E''$ . Nous écrivons la définition (27.816) de  $J$  pour  $x = J^{-1}(\xi)$  :

$$\varphi(J^{-1}(\xi)) = J(J^{-1}(\xi))\varphi = \xi(\varphi). \quad (27.1089)$$

Nous devons donc vérifier la continuité, pour chaque  $\varphi \in E'$  de l'application d'évaluation

$$\begin{aligned} ev_\varphi: (E', \sigma(E', E''))' &\rightarrow \mathbb{K} \\ \xi &\mapsto \xi(\varphi). \end{aligned} \quad (27.1090)$$

Par la définition 27.217(2) de la topologie faible sur  $E''$ , les applications  $ev_\varphi$  sont continues. Nous en concluons par la proposition 7.61 que l'application  $J^{-1}$  est continue.

Par ailleurs la proposition 27.180(2) dit que  $\overline{B_E(0, 1)} = J^{-1}(\overline{B_{E''}(0, 1)})$ , et le théorème de Banach-Alaoglu 27.234 appliqué à  $E'$  dit que  $\overline{B_{E''}(0, 1)}$  est compacte dans  $\sigma(E'', E')$ . Or l'image d'un compact par une application continue est compacte (proposition 7.274). Donc  $\overline{B_E(0, 1)}$  est compact.

- (2)  $\Leftarrow$  Nous supposons que  $\overline{B_E(0, 1)}$  est  $\sigma(E, E')$ -compacte. Nous savons par la proposition 27.182 que l'application

$$J: (E, \sigma(E, E')) \rightarrow (E'', \sigma(E'', E')) \quad (27.1091)$$

est continue. Comme l'image d'un compact est compacte, la partie  $J(\overline{B_E(0, 1)})$  est  $\sigma(E'', E')$ -compacte et donc fermée. Mais le théorème de Goldstine 27.232 dit que  $J(\overline{B_E(0, 1)})$  est  $\sigma(E'', E')$ -dense dans  $\overline{B_{E''}(0, 1)}$ . Donc

$$J(\overline{B_E(0, 1)}) = \overline{B_{E''}(0, 1)} \quad (27.1092)$$

par la proposition 7.25.

146. Je n'ai pas vérifié. Si ça pose un problème, écrivez-moi.

147. Théorème 7.8.

148. Définition 27.179.

149. Définition 27.217.



Nous en déduisons que  $J(E) = E''$ . En effet si  $\xi \in E''$ , alors il existe  $\lambda \in \mathbb{K}$  tel que  $\lambda\xi \in \overline{B_{E''}(0, 1)}$  (par exemple  $\lambda = 1/\|\xi\|$ ). Il existe donc  $x \in \overline{B_E(0, 1)}$  tel que  $J(x) = \lambda\xi$ . Et donc  $J(x/\lambda) = \xi$ . Cela prouve que  $\xi \in J(E)$ .

Nous avons démontré que  $J$  est surjective et donc que  $E$  est réflexif.

□

PROPooRNORooDjGqWy

**Proposition 27.236** ([220]).

Soient un espace de Banach réflexif  $E$  et un sous-espace vectoriel fermé  $M$ . Alors  $M$  est réflexif<sup>150</sup>.

*Démonstration.* Nous pouvons considérer deux topologies sur  $M$  :

- (1)  $\tau_1$  : la topologie induite de  $\sigma(E, E')$ ,
- (2)  $\tau_2$  : la topologie faible de  $M$  lui-même, c'est-à-dire  $\sigma(M, M')$ .

Et voici plein de petites affirmations successives.

- (1) Nous avons déjà vu dans la proposition 27.230 que  $\tau_1 = \tau_2$ .
- (2) Le théorème de Kakutani 27.235 dit que  $\overline{B_E(0, 1)}$  est  $\sigma(E, E')$ -compacte.
- (3) La partie  $M$  est convexe et  $\tau_E$ -fermée, donc  $M$  est  $\sigma(E, E')$ -fermée par la proposition 27.229.
- (4) Pour la topologie  $\sigma(E, E')$ , l'égalité  $\overline{B_M(0, 1)} = \overline{B_E(0, 1)} \cap M$  est une intersection d'un compact avec un fermé.  $\overline{B_M(0, 1)}$  est  $\sigma(E, E')$ -compacte par 7.122.
- (5) La partie  $\overline{B_M(0, 1)}$  est  $\sigma(E, E')$ -compacte et contenue dans  $M$ . La proposition 7.32 indique donc que  $\overline{B_M(0, 1)}$  est compacte pour la topologie de  $M$  induite de  $\sigma(E, E')$ , c'est-à-dire pour  $\tau_1$ .
- (6) Vu que  $\overline{B_M(0, 1)}$  est  $\tau_1$ -compacte, elle est aussi  $\tau_2$ -compacte parce que  $\tau_1 = \tau_2$ .
- (7) La boule  $\overline{B_M(0, 1)}$  est  $\sigma(M, M')$ -compacte. Donc le théorème de Kakutani 27.235 dit que  $M$  est réflexif.

□

PROPooBBNBooGcXDRH

**Proposition 27.237** ([220, 1]).

Un espace de Banach est réflexif si et seulement si son dual est réflexif.

*Démonstration.* En deux parties.

- (1)  $\Rightarrow$  Nous devons montrer que l'application  $J' : E' \rightarrow E'''$  est surjective. Soit  $\alpha \in E'''$  ; nous considérons

$$\begin{aligned} f_\alpha : E &\rightarrow \mathbb{K} \\ x &\mapsto (\alpha \circ J)(x). \end{aligned} \tag{27.1093}$$

En tant que composition d'applications linéaires continues, ce  $f_\alpha$  est un élément de  $E'$ . Nous montrons à présent que  $J'(f_\alpha) = \alpha$ . Soit  $\xi \in E''$ . Nous avons d'abord  $J'(f_\alpha)\xi = \xi(f_\alpha)$ . Mais comme  $J$  est surjective, il existe  $x \in E$  tel que  $\xi = J(x)$ . Pour ce  $x$  nous avons

$$J'(f_\alpha)\xi = \xi(f_\alpha) = J(x)f_\alpha = f_\alpha(x) = \alpha(J(x)) = \alpha(\xi). \tag{27.1094}$$

Nous avons donc bien  $J'(f_\alpha)\xi = \alpha(\xi)$  pour tout  $\xi \in E''$ .

- (2)  $\Leftarrow$  La partie  $\Rightarrow$  appliquée à  $E'$  montre que  $E''$  est réflexif, la proposition 27.181 dit que  $J(E)$  est fermé, et donc la proposition 27.236 dit que  $J(E)$  est réflexif.

L'application  $J : E \rightarrow J(E)$  est une bijection isométrique. Donc  $E$  est réflexif si et seulement si  $J(E)$  est réflexif par la proposition 27.184. Nous venons de prouver que  $J(E)$  est réflexif. Donc  $E$  est réflexif.

□

PROPooPVVYooMZjQsq

**Proposition 27.238** ([220]).

Une suite bornée dans un espace de Banach réflexif<sup>151</sup> contient une sous-suite faiblement convergente.

150. L'espace  $M$  est de Banach par la proposition 7.318

151. Définition 27.179.

### 27.17.6 Espaces $L^\infty$

**Lemme 27.239** ([220]).

LEMooMSYAooGEMgoc

Si  $(\Omega, \mathcal{A}, \mu)$  est un espace mesuré  $\sigma$ -fini, l'espace  $L^1(\Omega, \mathcal{A}, \mu)$  n'est pas réflexif.

**Lemme 27.240** ([220]).

LEMooUSXTooFvpsVd

Si  $(\Omega, \mathcal{A}, \mu)$  est un espace mesuré  $\sigma$ -fini, alors  $L^\infty(\Omega, \mathcal{A}, \mu)$  n'est pas réflexif<sup>152</sup>.

*Démonstration.* Supposons que  $L^\infty(\Omega, \mathcal{A}, \mu)$  est réflexif. Le théorème de représentation de Riesz 27.187(2) dit que  $(L^1)'$  est en bijection linéaire isométrique avec  $L^\infty$ ; le lemme 27.184 dit alors que  $(L^1)'$  est réflexif<sup>153</sup>. La proposition 27.237 dit alors que  $L^1$  est réflexif.

Or le lemme 27.239 dit que  $L^1$  n'est pas réflexif.  $\square$

La proposition suivante est souvent présentée en disant que l'inclusion  $L^1 \subset (L^\infty)'$  est stricte, ou, pire, en disant que  $(L^\infty)'$  est strictement plus grand que  $L^1$ . Cette façon de dire est un gros abus de langage. D'abord  $L^1$  n'est même pas inclus dans  $L^\infty$ ; ce sont deux ensembles qui n'ont rien à voir. Ensuite, ce que signifie réellement cette proposition est seulement que la première injection  $L^1 \rightarrow (L^\infty)'$  qui nous tombe sous la main (celle du théorème de représentation de Riesz) n'est pas surjective<sup>154</sup>.

PROPooXXRQooNSBZOi

**Proposition 27.241.**

Soit  $g \in L^1(\mathbb{R}^d)$ . L'application

$$\begin{aligned} \Phi_g: L^\infty(\mathbb{R}^d) &\rightarrow \mathbb{C} \\ f &\mapsto \int_{\mathbb{R}^d} f \bar{g} \end{aligned} \quad (27.1095)$$

est linéaire et bien définie.

L'application

$$\begin{aligned} \Phi: L^1(\mathbb{R}^d) &\rightarrow L^\infty(\mathbb{R}^d)' \\ g &\mapsto \Phi_g \end{aligned} \quad (27.1096)$$

n'est pas surjective.

*Démonstration.* Prouvons d'abord que  $\Phi_g$  est bien définie. Vu que  $f \in L^\infty$ , il existe  $M \in \mathbb{R}$  tel que  $|f(x)| < M$  sur  $\mathbb{R}^d \setminus A$  où  $A$  est de mesure nulle dans  $\mathbb{R}^d$ . La fonction  $x \mapsto |f(x)g(x)|$  est donc majorée par la fonction intégrable  $x \mapsto M|g(x)|$  qui est intégrable (sur  $\mathbb{R}^d \setminus A$ ). L'intégrabilité de  $f\bar{g}$  n'est donc pas un problème.

Le fait que  $\Phi$  prenne ses valeurs dans  $(L^\infty)'$  est le théorème 27.188(2b).

Le fait qu'elle ne soit pas surjective est la proposition 27.192.  $\square$

## 27.18 Espace de Schwartz

Pour un multiindice  $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$ , nous notons

$$\partial^\alpha \varphi = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} \varphi \quad (27.1097)$$

pour peu que la fonction  $\varphi$  soit  $|\alpha| = \alpha_1 + \dots + \alpha_d$  fois dérivable.

DefHHyQooK

**Définition 27.242.**

Soit  $\Omega \subset \mathbb{R}^d$ . L'espace de Schwartz  $\mathcal{S}(\Omega)$  est le sous-ensemble de  $C^\infty(\Omega)$  des fonctions dont toutes les dérivées décroissent plus vite que tout polynôme :

$$\mathcal{S}(\Omega) = \{ \varphi \in C^\infty(\Omega) \text{ tel que } \forall \alpha, \beta \in \mathbb{N}^d, p_{\alpha, \beta}(\varphi) < \infty \} \quad (27.1098)$$

<sup>152</sup>. Définition 27.179.

<sup>153</sup>. Dans de nombreuses références, par exemple [220], il est simplement dit que  $(L^1)' = L^\infty$ . C'est un abus de notation qui permet de se passer du lemme 27.184.

<sup>154</sup>. Si vous savez comment prouver qu'il n'existe pas de surject de  $L^1$  vers  $(L^\infty)'$ , écrivez-moi.

où nous avons considéré

$$p_{\alpha,\beta}(\varphi) = \sup_{x \in \Omega} |x^\beta (\partial^\alpha \varphi)(x)| = \|x^\beta \partial^\alpha \varphi\|_\infty. \quad (27.1099) \quad \text{EqQWdChCv}$$

Pour simplifier les notations (surtout du côté de Fourier), nous allons parfois écrire  $M_i \varphi$  pour la fonction  $x \mapsto x_i \varphi(x)$ .

**Exemple 27.243.**

La fonction  $e^{-x^2}$  est une fonction à décroissance rapide sur  $\mathbb{R}$ . △

**Définition 27.244.**

Une fonction  $f: \mathbb{R}^d \rightarrow \mathbb{C}$  est dite à **décroissance rapide** si elle décroît plus vite que n'importe quel polynôme. Plus précisément, si pour tout polynôme  $Q$ , il existe un  $r > 0$  tel que  $|f(x)| < \frac{1}{|Q(x)|}$  pour tout  $\|x\| \geq r$ .

PropCSmzwGv

**Proposition 27.245.**

Une fonction Schwartz est à décroissance rapide.

*Démonstration.* Nous commençons par considérer un polynôme  $P$  donné par

$$P(x) = \sum_k c_k x^{\beta_k} \quad (27.1100)$$

où les  $\beta_k$  sont des multiindices, les  $c_k$  sont des constantes et la somme est finie. Nous avons la majoration

$$\sup_{x \in \mathbb{R}^d} |\varphi(x) P(x)| \leq \sum_k \sup_x |c_k \varphi(x) x^{\beta_k}| \leq \sum_k |c_k| p_{0,\beta_k}(\varphi) < \infty. \quad (27.1101)$$

Nous allons noter  $M_P$  la constante  $\sum_k |c_k| p_{0,\beta_k}(\varphi)$ , de sorte que pour tout  $x \in \mathbb{R}^d$  nous avons  $|\varphi(x) P(x)| \leq M_P$  et donc

$$|\varphi(x)| \leq \frac{M_P}{|P(x)|} = \frac{1}{|\frac{1}{M_P} P(x)|}. \quad (27.1102)$$

Notons que cette inégalité est a fortiori correcte pour les  $x$  sur lesquels  $P$  s'annule.

Soit maintenant un polynôme  $Q$ . Nous considérons le polynôme  $P(x) = \|x\| Q(x)$ . Étant de plus haut degré, pour toute constante  $C$  il existe un rayon  $r_C$  tel que  $|P(x)| \geq C|Q(x)|$  pour tout  $|x| \geq r_C$ . En particulier pour  $|x| \geq r_{M_P}$  nous avons

$$|P(x)| \geq M_P |Q(x)| \quad (27.1103)$$

et donc, pour ces  $x$ ,

$$|\varphi(x)| \leq \frac{1}{|\frac{1}{M_P} P(x)|} \leq \frac{1}{|Q(x)|}. \quad (27.1104)$$

La première inégalité est valable pour tout  $x$ , et la seconde pour  $\|x\| \geq r_{M_P}$ . □

CORooZFPSooHCFUSH

**Corolaire 27.246 ([1]).**

Soit  $\varphi$  une fonction Schwartz sur  $\mathbb{R}^m \times \mathbb{R}^n$ . Alors la fonction

$$y \mapsto \sup_{x \in \mathbb{R}^n} |\varphi(x, y)| \quad (27.1105)$$

est intégrable.

*Démonstration.* Soit un polynôme  $Q$  en la variable  $y$ . Par la proposition 27.245, il existe  $r > 0$  tel que

$$|\varphi(x, y)| < \frac{1}{Q(y)} \quad (27.1106)$$

pour tout  $\|(x, y)\| > r$ . A fortiori l'inégalité tient pour tout  $|y| > r$ . Donc

$$\int_{\mathbb{R}^m} \sup_{x \in \mathbb{R}^n} |\varphi(x, y)| dy = \int_{\|y\| \leq r} \sup_x |\varphi(x, y)| dy + \int_{\|y\| > r} \sup_x |\varphi(x, y)| dy. \quad (27.1107)$$

La première intégrale est bornée par  $\text{Vol}(B(0, r))\|\varphi\|_\infty$  tandis que la seconde est bornée par l'intégrale de  $\frac{1}{Q(y)}$ . En prenant  $Q$  de degré suffisamment élevé en toutes les composantes de  $y$  nous avons intégrabilité.  $\square$

### 27.18.1 Topologie

#### Lemme-Définition 27.247.

Les  $p_{\alpha,\beta}$  donnés par l'équation (27.1099) ci-dessus sont des seminormes<sup>155</sup>. La topologie considérée sur  $\mathcal{S}(\mathbb{R}^d)$  est celle des seminormes  $p_{\alpha,\beta}$ .

#### 27.248.

Nous avons un enchainement de résultats qui nous aident à prouver la continuité d'une application  $T: \mathcal{S}(\mathbb{R}^d) \rightarrow X$ .

- (1) La topologie de  $\mathcal{S}(\mathbb{R}^d)$  est donnée par une famille dénombrable de seminormes. Donc la proposition 7.402 nous dit que  $\mathcal{S}(\mathbb{R}^d)$  est métrisable.
- (2) La proposition 7.349 nous dit alors que si  $X$  est métrique, toute application séquentiellement continue  $T: \mathcal{S}(\mathbb{R}^d) \rightarrow X$  est continue.
- (3) Donc si  $X$  est métrique, il suffit de prouver que pour  $f_n \xrightarrow{\mathcal{S}(\mathbb{R}^d)} 0$  nous avons  $T(f_n) \xrightarrow{X} 0$  où  $f_n: \mathcal{S}(\mathbb{R}^d) \rightarrow X$ . Dans les cas usuels,  $T$  sera une distribution et  $X = \mathbb{C}$ .
- (4) En vertu de la proposition 7.398, la convergence  $f_n \xrightarrow{\mathcal{S}(\mathbb{R}^d)} 0$  signifie que pour tout choix de multiindice  $\alpha$  et  $\beta$ ,  $p_{\alpha,\beta}(f_n) \rightarrow 0$ , c'est-à-dire

$$\|x^\beta \partial^\alpha f_n\|_\infty \rightarrow 0. \quad (27.1108)$$

- (5) Et enfin, la technique pour montrer que  $T: \mathcal{S}(\mathbb{R}^d) \rightarrow \mathbb{C}$  est continue est de montrer que sous l'hypothèse d'avoir (27.1108) pour tout choix de  $\alpha$  et  $\beta$ , nous avons  $T(f_n) \rightarrow 0$  dans  $\mathbb{C}$ .

#### Lemme 27.249 ([774]).

La topologie sur  $\mathcal{S}(\mathbb{R}^d)$  est donnée aussi par les seminormes

$$q_{n,m} = \max_{|\alpha| \leq n} \sup_{x \in \mathbb{R}^d} (1 + \|x\|)^m |\partial^\alpha \varphi(x)|. \quad (27.1109)$$

Autrement dit, une suite  $\varphi_n \xrightarrow{\mathcal{S}(\mathbb{R}^d)} 0$  si et seulement si  $q_{n,m}(\varphi) \rightarrow 0$  pour tout  $n$  et  $m$ .

Le fait que les  $q_{n,m}(\varphi)$  restent bornés est la proposition 27.245. Cependant ce lemme est plus précis parce qu'en disant seulement que  $\varphi$  est majoré par des polynôme, nous ne disons pas que les polynômes correspondants aux  $\varphi_n$  tendent vers zéro si  $\varphi_n \xrightarrow{\mathcal{S}} 0$ . Et d'ailleurs on ne sait pas très bien ce que signifierait  $P_n \rightarrow 0$  pour une suite de polynômes.

#### Proposition 27.250.

Pour  $p \in [1, \infty]$ , l'espace  $\mathcal{S}(\mathbb{R}^d)$  s'injecte continument dans  $L^p(\mathbb{R}^d)$ .

*Démonstration.* L'injection dont nous parlons est l'identité ou plus précisément l'identité suivie de la prise de classe. Il faut vérifier que cela est correct et continu, c'est-à-dire d'abord qu'une fonction à décroissance rapide est bien dans  $L^p$  et ensuite que si  $f_n \xrightarrow{\mathcal{S}} 0$ , alors  $f_n \xrightarrow{L^p} 0$ .

Commençons par  $p = \infty$ . Alors  $\|f_n\|_\infty = p_{0,0}(f_n) \rightarrow 0$  parce que si  $f_n \xrightarrow{\mathcal{S}} 0$ , alors en particulier  $p_{0,0}(f_n) \rightarrow 0$ .

Au tour de  $p < \infty$  maintenant. Nous savons qu'en dimension  $d$ , la fonction

$$x \mapsto \frac{1}{(1 + \|x\|)^s} \quad (27.1110)$$

155. Définition 7.380.

est intégrable dès que  $s > d$ . Pour toute valeur de  $m$  nous avons

$$\|\varphi\|_p^p = \int_{\mathbb{R}^d} |\varphi(x)|^p dx = \int_{\mathbb{R}^d} \frac{|(1 + \|x\|)^m \varphi(x)|^p}{(1 + \|x\|)^{mp}} \leq \int_{\mathbb{R}^d} \frac{q_{0,m}(\varphi)^p}{(1 + \|x\|)^{mp}}. \quad (27.1111)$$

En choisissant  $m$  de telle sorte que  $mp > d$ , nous avons convergence de l'intégrale et donc  $\|\varphi\|_p < \infty$ . Nous retenons que

$$\|\varphi\|_p^p \leq C q_{0,m}(\varphi)^p \quad (27.1112)$$

pour une certaine constante  $C$  et un bon choix de  $m$ .

Ceci prouve que  $\mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ . Nous devons encore vérifier que l'inclusion est continue. Si  $\varphi_n \xrightarrow{\mathcal{S}} 0$ , alors en particulier nous avons  $q_{0,m}(\varphi_n) \rightarrow 0$  par le lemme 27.249. Par conséquent la majoration (27.1112) nous dit que  $\|\varphi_n\|_p \rightarrow 0$  également.  $\square$

En résumé, si  $\varphi_n \xrightarrow{\mathcal{S}(\mathbb{R}^d)} \varphi$  alors  $\varphi_n \xrightarrow{L^p} \varphi$ .

**Théorème 27.251** ([523]).

Soit  $\mu$  une mesure sur les boréliens de  $\mathbb{R}^n$  finie sur les compacts. Alors  $\mathcal{D}(\mathbb{R}^n)$  est dense dans  $L^1(\mathbb{R}^n, \text{Bor}(\mathbb{R}^n), \mu)$ .

**Proposition 27.252** ([775]).

La partie  $\mathcal{D}(\mathbb{R}^d)$  est dense dans  $\mathcal{S}(\mathbb{R}^d)$ .

*Démonstration.* Soit  $f \in \mathcal{S}(\mathbb{R}^d)$ , et  $\phi$ , une fonction de  $\mathcal{D}(\mathbb{R}^d)$  telle que  $\phi(x) = 1$  pour  $|x| \leq 1$  (l'existence de telles fonctions est discutée en 15.14.1). Soit aussi  $\phi_k(x) = \phi(x/k)$ . Nous posons

$$f_k(x) = \phi_k(x)f(x), \quad (27.1113)$$

et nous allons prouver que pour tout multiindice  $\alpha$  et  $\gamma$ ,

$$p_{\alpha,\gamma}(f_k - f) = \|x^\gamma \partial^\alpha (f_k - f)\|_\infty \rightarrow 0. \quad (27.1114)$$

Pour cela nous allons noter  $\beta \leq \alpha$  lorsque  $\beta$  est un multiindice contenu dans  $\alpha$ . En utilisant la dérivée du produit nous avons

$$(\partial^\alpha f_k)(x) = \sum_{\beta \leq \alpha} (\partial^{\alpha-\beta} \phi_k)(x) \partial^\beta f(x) \quad (27.1115a)$$

$$= \sum_{\beta \leq \alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta} \phi)(x/k) (\partial^\beta f)(x) \quad (27.1115b)$$

$$= \sum_{\beta < \alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta} \phi)(x/k) (\partial^\beta f)(x) + \phi(x/k) (\partial^\alpha f)(x). \quad (27.1115c)$$

Nous devons donc étudier et majorer

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} |x^\gamma \partial^\alpha (f_k - f)| &\leq \sup_{x \in \mathbb{R}^d} |x^\gamma \sum_{\beta < \alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta} \phi)(x/k) (\partial^\beta f)(x)| \\ &\quad + \sup_{x \in \mathbb{R}^d} |x^\gamma (\phi(x/k) - 1) (\partial^\alpha f)(x)| \end{aligned} \quad (27.1116)$$

En ce qui concerne le second terme, soit  $\epsilon > 0$ , vu que  $f$  est Schwartz, il existe  $R$  tel que

$$|x^\gamma (\partial^\alpha f)(x)| < \epsilon \quad (27.1117)$$

dès que  $\|x\| > R$ . En prenant  $k > R$ ,

$$|x^\gamma (\partial^\alpha f)(x)| \begin{cases} = 0 & \text{si } \|x\| < R \\ \leq \epsilon & \text{si } \|x\| > R. \end{cases} \quad (27.1118)$$

En ce qui concerne le premier terme,

$$\sup_{x \in \mathbb{R}^d} |x^\gamma \sum_{\beta < \alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta} \phi)(x/k) (\partial^\beta f)(x)| \quad (27.1119a)$$

$$\leq \frac{1}{k} \sup_x \left| \sum_{\beta < \alpha} (\partial^{\alpha-\beta} \phi)(x/k) (x^\gamma \partial^\beta f)(x) \right| \quad (27.1119b)$$

$$= \frac{1}{k} \sup_x \left| \sum_{\beta < \alpha} (\partial^{\alpha-\beta} \phi)(x/k) p_{\beta,\gamma}(f) \right| \quad (27.1119c)$$

La somme ne contient qu'un nombre fini de  $\beta$  différents, donc nous pouvons considérer un nombre  $K$  qui majore tous les  $p_{\beta,\gamma}(f)$  en même temps. La partie avec  $\phi$  peut être majorée par  $\|\partial^{\alpha-\beta} \phi\|_\infty$  (qui est fini) dont nous pouvons prendre le maximum sur  $\beta < \alpha$ . Toute l'expression dans la somme est donc majorée par un nombre qui ne dépend ni de  $x$  ni de  $\beta$ . Vu que la somme est finie, elle est majorée par ce nombre multiplié par le nombre de termes dans la somme et au final

$$\sup_{x \in \mathbb{R}^d} |x^\gamma \sum_{\beta < \alpha} k^{-|\alpha-\beta|} (\partial^{\alpha-\beta} \phi)(x/k) (\partial^\beta f)(x)| \leq \frac{K'}{k}. \quad (27.1120)$$

La limite  $k \rightarrow \infty$  ne fait alors plus de doute.  $\square$

### Remarque 27.253.

Vu la topologie de  $\mathcal{S}(\mathbb{R}^d)$  (définition 27.247), la convergence  $f_k \xrightarrow{\mathcal{S}(\mathbb{R}^d)} f$  peut être exprimée par le fait que pour tout  $k, l$ ,

$$t^k f_n^{(l)} \xrightarrow{\text{unif}} t^k f^{(l)}. \quad (27.1121)$$

C'est-à-dire convergence uniforme de toutes les dérivées multipliées par n'importe quel polynôme.

## 27.18.2 Produit de convolution

PROPooUNFYooYdbSbJ

**Proposition 27.254** (Stabilité de Schwartz par convolution<sup>156</sup> [745]).

Si  $\varphi \in L^1(\mathbb{R}^d)$  et  $\psi \in \mathcal{S}(\mathbb{R}^d)$ , alors  $\varphi * \psi \in \mathcal{S}(\mathbb{R}^d)$ .

*Démonstration.* Nous devons prouver que

$$p_{\alpha,\beta}(\varphi * \psi) = \sup_{x \in \mathbb{R}^d} |x^\beta (\partial^\alpha (\varphi * \psi))(x)| \quad (27.1122)$$

est borné pour tout multiindices  $\alpha$  et  $\beta$ . En appliquant  $|\alpha|$  fois la proposition 27.84, nous mettons toutes les dérivées sur  $\psi$  :  $\partial^\alpha (\varphi * \psi) = (\varphi * \partial^\alpha \psi)$ . Cela étant fait, nous majorons

$$|x^\beta (\varphi * \partial^\alpha \psi)(x)| \leq |x^\beta| \int_{\mathbb{R}^d} |\varphi(y)| \underbrace{|(\partial^\alpha \psi)(x-y)|}_{\leq \|\partial^\alpha \psi\|_\infty} dy \quad (27.1123a)$$

$$\leq |x^\beta| \|\partial^\alpha \psi\|_\infty \int_{\mathbb{R}^d} |\varphi(y)| dy \quad (27.1123b)$$

$$\leq p_{\alpha,\beta}(\psi) \|\varphi\|_{L^1}. \quad (27.1123c)$$

Par conséquent,  $p_{\alpha,\beta}(\varphi * \psi) \leq \|\varphi\|_{L^1} p_{\alpha,\beta}(\psi) < \infty$ .  $\square$

## 27.19 Théorème de Montel

ThoXLyCzol

**Théorème 27.255** (Montel[126]).

Soient  $\Omega$  un ouvert de  $\mathbb{C}$  et  $\mathcal{F}$  une famille de fonctions holomorphes sur  $\Omega$ , uniformément bornée sur tout compact de  $\Omega$ . Alors de toute suite dans  $\mathcal{F}$  nous pouvons extraire une sous-suite convergent uniformément sur tout compact de  $\Omega$ .

156. Définition 27.78.

**Démonstration.** (1) Un ensemble équicontinu

Nous commençons par prendre une suite de compacts dans  $\Omega$  comme dans le lemme 7.358, et une suite  $\delta_n$  de réels strictement positifs tels que

$$B(z, 2\delta_n) \subset K_{n+1} \quad (27.1124)$$

pour tout  $z \in K_n$ . Soient  $x, y \in K_n$  tels que  $|x - y| < \delta_n$ ; nous notons  $\partial B(x, 2\delta_n)$  le cercle de rayon  $2\delta_n$  autour de  $x$ , parcouru dans le sens positif. La formule de Cauchy 26.217 nous donne

$$f(x) - f(y) = \frac{1}{2\pi i} \int_{\partial B} \left( \frac{f(\xi)}{\xi - x} - \frac{f(\xi)}{\xi - y} \right) d\xi = \frac{x - y}{2\pi i} \int_{\partial B} \frac{f(\xi)}{(\xi - x)(\xi - y)} d\xi \quad (27.1125)$$

Nous majorons ça par

$$|f(x) - f(y)| \leq \frac{|x - y|}{2\pi} \int_{\partial B} \frac{|f(\xi)|}{2\delta_n^2} d\xi \leq \frac{|x - y|}{\delta_n} M_n. \quad (27.1126)$$

Justifications :

- $|\xi - x| = 2\delta_n$  et  $|\xi - y| \geq \delta_n$  parce que  $\xi$  est au mieux sur le rayon passant par  $x$  et  $y$ .
- $|f(\xi)| \leq M_n$  où  $M_n$  est la borne uniforme de  $\mathcal{F}$  sur le compact  $K_n$ .
- Nous avons aussi fini par calculer l'intégrale dans laquelle il ne restait plus rien, ça a donné la circonférence du cercle de rayon  $2\delta_n$ .

Jusqu'à présent nous avons prouvé que l'ensemble

$$\mathcal{F}_n = \{f|_{K_n} \text{ tel que } f \in \mathcal{F}\} \quad (27.1127)$$

est équicontinu. Il est aussi équiborné par hypothèse.

(2) Application du théorème d'Ascoli

L'ensemble  $\mathcal{F}_n$  vérifie les hypothèses du théorème d'Ascoli 27.7. Donc l'ensemble  $\mathcal{F}_n$  est relativement compact dans  $C(K_n, \mathbb{C})$  pour la norme uniforme. Autrement dit l'ensemble  $\bar{\mathcal{F}}$  est compact et si nous avons une suite de fonctions dans  $\mathcal{F}_n$ , il existe une sous-suite convergeant dans  $\bar{\mathcal{F}}$ , c'est-à-dire uniformément. Autrement dit il existe une fonction strictement croissante  $\varphi: \mathbb{N} \rightarrow \mathbb{N}$  telle que la suite  $k \mapsto f_{\varphi(k)}$  converge uniformément sur  $K_n$ . La limite n'est cependant pas spécialement dans  $\mathcal{F}_n$ .

(3) L'argument diagonal

La suite  $k \mapsto f_{\varphi_1 \circ \dots \circ \varphi_k(k)}$  converge uniformément sur tous les  $K_n$ . Si  $K$  est un compact de  $\Omega$ , alors les petites propriétés sympas du lemme 7.358 nous disent que  $K \subset \text{Int}(K_m)$  pour un certain  $m$ . Ladite suite convergeant uniformément sur  $K_m$ , elle converge uniformément sur  $K$  et nous avons montré la convergence uniforme sur tout compact de  $\Omega$ . □

**Corolaire 27.256** ([126]).

Soient  $\Omega$  un ouvert connexe borné de  $\mathbb{C}$  et  $a \in \Omega$ . Soit  $f$  holomorphe sur  $\Omega$  telle que  $f(a) = a$  et  $|f'(a)| < 1$ .

Alors de  $(f^n)$  on peut extraire une sous-suite convergeant uniformément sur tout compact de  $\Omega$  vers la fonction constante  $a$ .

*Démonstration.* Nous considérons un voisinage de  $a$  inclus dans  $\Omega$ ; sachant que  $|f(a)| < 1$ , nous trouvons un voisinage encore plus petit de  $a$  sur lequel  $|f'(z)| < 1$ . Soit donc  $r$  tel que  $\bar{B}(a, r) \subset \Omega$  et tel que  $|f'(z)| < 1$  sur  $\bar{B}(a, r)$ . Étant donné que  $f'(z)$  est continue sur le compact  $\bar{B}(a, r)$ , nous en prenons le maximum  $\lambda$  (qui est strictement inférieur à 1) et nous avons au final

$$|f'(z)| \leq \lambda < 1 \quad (27.1128)$$



pour tout  $z \in \overline{B(a, r)}$ . Le théorème des accroissements finis 12.340 nous dit que

$$|f(z) - a| \leq \lambda |z - a| \quad (27.1129)$$

pour tout  $z \in \overline{B(a, r)}$ . C'est ici que nous utilisons l'hypothèse de convexité de  $\Omega$ . Nous montrons alors par récurrence que

$$|f^n(z) - a| \leq \lambda^n |z - a| \leq \lambda^n r \leq r. \quad (27.1130)$$

L'ensemble  $A = \{f^n \text{ tel que } n \geq 1\}$  est donc uniformément borné sur  $\overline{B(a, r)}$  par  $a + r$ . Autre manière de le dire : pour tout  $z \in \overline{B(a, r)}$  nous avons

$$f^n(z) \in \overline{B(a, r)}. \quad (27.1131)$$

La suite  $(f^n)$  est donc uniformément bornée sur tout compact de  $B(a, r)$ . Le théorème de Montel 27.255 nous indique que l'on peut extraire une sous-suite convergente uniformément sur tout compact. Au vu de (27.1130) cette convergence ne peut avoir lieu que vers une fonction  $g$  qui vaut la constante  $a$  sur  $B(a, r)$ .

D'autre part la fonction  $g$  est holomorphe en tant que limite uniforme de fonctions holomorphes, théorème 26.109. Or une fonction holomorphe constante sur un ouvert est constante sur tout son domaine d'holomorphie (principe d'extension analytique, théorème 17.138).  $\square$

## 27.20 Espaces de Bergman

Source : [555].

Soit  $\Omega$  un borné dans  $\mathbb{C}$  et  $D$  le disque unité ouvert de  $\mathbb{C}$ .

### Définition 27.257.

*L'espace de Bergman* sur  $\Omega$ , noté  $A^2(\Omega)$  est l'espace des fonctions holomorphes sur  $\Omega$  qui sont en même temps dans  $L^2(\Omega)$ .

Nous mettons sur  $A^2(\Omega)$  le produit scalaire usuel hérité de  $L^2$  :

$$\langle f, g \rangle = \int_{\Omega} f(z) \overline{g(z)} dz. \quad (27.1132)$$

### Lemme 27.258.

Soient un compact  $K \subset \Omega$  et une fonction  $f \in A^2(\Omega)$ . Alors

$$\max_{z \in K} |f(z)| \leq \frac{1}{\sqrt{\pi}} \frac{1}{d(K, \partial\Omega)} \|f\|_2. \quad (27.1133)$$

*Démonstration.* Soient  $a \in \Omega$  et  $r > 0$  tels que  $B(a, r) \subset \Omega$ . Nous considérons aussi  $\rho \leq r$ . La formule de Cauchy (26.217) nous donne

$$f(a) = \frac{1}{2\pi i} \int_{B(a, \rho)} \frac{f(\xi)}{\xi - a} d\xi = \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho e^{i\theta}) d\theta \quad (27.1134)$$

où nous avons utilisé le chemin  $\gamma(\theta) = a + \rho e^{i\theta}$ ,  $\gamma'(\theta) = i\rho e^{i\theta}$  et  $\rho = |\xi - a|$ . Maintenant une astuce est d'écrire

$$\frac{r^2}{2} f(a) = \int_0^r f(a) \rho d\rho, \quad (27.1135)$$

et d'y substituer la valeur de  $f(a)$  que nous venons de calculer :

$$\frac{r^2}{2} f(a) = \int_0^r \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho e^{i\theta}) d\theta \rho d\rho \quad (27.1136a)$$

$$= \frac{1}{2\pi} \int_{B(a, r)} f(z) dz \quad \text{passage aux polaires} \quad (27.1136b)$$

$$= \frac{1}{2\pi} \langle 1, f \rangle_B \quad \text{produit scalaire sur } B(a, r) \quad (27.1136c)$$

$$\leq \frac{1}{2\pi} \sqrt{\langle 1, 1 \rangle_B \langle f, f \rangle_B} \quad (27.1136d)$$



Nous avons donc

$$r^2 f(a) \leq \frac{1}{\pi} \sqrt{\langle 1, 1 \rangle_B \langle f, f \rangle_B}, \quad (27.1137)$$

et donc

$$\pi r^2 f(a) \leq \sqrt{\pi r^2} \|f\|_2, \quad (27.1138)$$

parce que  $\langle f, f \rangle_B \leq \|f\|_2^2$ . En effet le produit scalaire  $\|\cdot\|_2$  est donné par une intégrale sur  $\Omega$  alors que  $B(a, r) \subset \Omega$  et que la fonction qu'on y intègre est positive (c'est  $|f(z)|^2$ ). En simplifiant,

$$f(a) \leq \frac{1}{\sqrt{\pi} r} \|f\|_2. \quad (27.1139)$$

Mais  $r$  a été choisi pour avoir  $B(a, r) \subset \Omega$ , donc  $r \leq d(a, \partial\Omega)$  et

$$|f(a)| \leq \frac{1}{d(a, \partial\Omega) \sqrt{\pi}} \|f\|_2. \quad (27.1140)$$

Maintenant si nous prenons  $a \in K$ , nous avons encore la minoration  $d(a, \partial K) \leq d(a, \partial\Omega)$  et donc

$$|f(a)| \leq \frac{1}{d(a, \partial K) \sqrt{\pi}} \|f\|_2. \quad (27.1141)$$

□

### **Théorème 27.259.**

Soit  $\Omega$  un ouvert de  $\mathbb{C}$ .

- (1) L'espace  $A^2(\Omega)$  est un espace de Hilbert.
- (2) Si  $D$  est la boule unité dans  $\mathbb{C}$ , une base hilbertienne de  $A^2(D)$  est donnée par les fonctions

$$e_n(z) = \sqrt{\frac{n+1}{\pi}} z^n \quad (27.1142)$$

pour  $n \geq 0$ .

*Démonstration.* Nous commençons par montrer que  $A^2(\Omega)$  est complet. Pour cela nous considérons une suite de Cauchy  $(f_n)$  dans  $A^2(\Omega)$  et un compact  $K \subset \Omega$ . Nous savons par le lemme 27.258 que

$$\max_{z \in K} |f_n(z) - f_m(z)| \leq \frac{1}{\sqrt{\pi} d(K, \partial\Omega)} \|f_n - f_m\|_2. \quad (27.1143)$$

Donc  $f_n$  converge uniformément sur  $K$ . Par le théorème de Weierstrass 26.109, la fonction  $f$  est holomorphe. Il existe donc une fonction holomorphe  $f$  qui est limite uniforme sur tout compact de  $\Omega$  de la suite  $(f_n)$ .

Mais  $L^2(\Omega)$  étant complet, la suite  $(f_n)$  a une limite  $g \in L^2(\Omega)$ . Ce que nous voudrions faire est prouver que  $f = g$ . Notons que tel quel, ce n'est pas vrai parce que  $f$  est une vraie fonction alors que  $g$  est une classe. Ce que nous enseigne la proposition 27.23 est qu'il existe une sous-suite (qu'on note  $(g_n)$ ) qui converge vers  $g$  presque partout. Dans cette dernière phrase,  $g_n$  et  $g$  sont de vraies fonctions, des représentants des classes dans  $L^2$ .

Nous déduisons que  $f = g$  presque partout (ici  $f$  et  $g$  sont les fonctions) parce que la sous-suite converge uniformément vers  $f$  en même temps que presque partout vers  $g$ . Donc  $f = g$  dans  $L^2(\Omega)$  (ici  $f$  et  $g$  sont les classes). Donc  $f \in L^2(\Omega)$  et l'espace  $A^2(\Omega)$  est de Hilbert.

Il nous faut encore prouver que  $(e_n)_{n \geq 0}$  est une base orthonormale. En ce qui concerne les produits scalaires,

$$\langle e_m, e_n \rangle = \sqrt{\frac{(m+1)(n+1)}{\pi}} \int_D z^n \bar{z}^m dz \quad (27.1144a)$$

$$= \sqrt{\frac{(m+1)(n+1)}{\pi^2}} \int_0^1 \rho d\rho \int_0^{2\pi} d\theta \rho^{m+n} e^{i\theta(n-m)} \quad (27.1144b)$$

$$= \sqrt{\frac{(m+1)(n+1)}{\pi^2}} \frac{1}{m+n+2} \underbrace{\int_0^{2\pi} e^{i\theta(n-m)} d\theta}_{2\pi\delta_{mn}} \quad (27.1144c)$$

$$= \sqrt{\frac{(n+1)^2}{\pi^2}} \frac{1}{2n+2} 2\pi\delta_{nm} \quad (27.1144d)$$

$$= \delta_{nm}. \quad (27.1144e)$$

Donc les fonctions données sont bien orthonormales. Nous devons montrer qu'elles sont denses dans  $A^2(D)$ . Soit  $f \in A^2(D)$  et  $c_n(f) = \langle f, e_n \rangle$ ; nous allons montrer que

$$\|f\|_2^2 = \sum_{n=0}^{\infty} |\langle f, e_n \rangle|^2, \quad (27.1145)$$

parce que le point (5) du théorème 25.46 nous indique que ce sera suffisant pour avoir une base hilbertienne.

Étant donné que  $f$  est holomorphe sur  $D$ , le théorème 26.60 nous développe  $f$  en série entière :

$$f(z) = \sum_{k=0}^{\infty} a_k z^k. \quad (27.1146) \quad \text{EqObkbPK}$$

En permutant la somme avec le produit scalaire,

$$c_n(f) = \int_D f(z) \bar{e}_n(z) = \sqrt{\frac{n+1}{\pi}} \int_D f(z) \bar{z}^n dz. \quad (27.1147)$$

Afin de profiter de la convergence uniforme de la série (27.1146) à l'intérieur de  $D$ , nous allons exprimer l'intégrale sur  $D$  comme une intégrale sur  $|z| < r$  en faisant tendre  $r$  vers 1 (par le bas). Pour ce faire nous considérons les fonctions

$$g_k(z) = \begin{cases} f(z) \bar{z}^n & \text{si } |z| < 1 - 1/k \\ 0 & \text{sinon.} \end{cases} \quad (27.1148)$$

Ces fonctions sont intégrables sur  $D$  et dominées par  $f(z) \bar{z}^n$  qui est intégrable sans dépendre de  $k$ . Mais nous avons évidemment  $g_k(z) \rightarrow f(z) \bar{z}^n$ . Le théorème de la convergence dominée permet alors de permuter l'intégrale et la limite  $k \rightarrow \infty$ . Cela nous permet d'écrire

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r \rightarrow 1^-} \int_{|z| < r} \bar{z}^n f(z) dz = \sqrt{\frac{n+1}{\pi}} \lim_{r \rightarrow 1^-} \int_{|z| < r} \sum_{k=0}^{\infty} a_k z^k \bar{z}^n. \quad (27.1149)$$

Par la convergence uniforme de la série entière à l'intérieur du disque  $D$  nous pouvons permuter l'intégrale et la somme (proposition 15.55) :

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r \rightarrow 1^-} \sum_{k=0}^{\infty} a_k \int_{|z| < r} z^k \bar{z}^n dz. \quad (27.1150)$$

L'intégrale proprement dite est vite calculée et vaut

$$\int_{|z| < 1} \bar{z}^n z^k dz = \frac{\pi r^{2n+2}}{n+1} \delta_{kn}. \quad (27.1151)$$

Nous pouvons donc continuer le calcul de  $c_n(f)$  en effectuant la somme sur  $k$  qui se réduit à changer  $k$  en  $n$  puis en effectuant la limite :

$$c_n(f) = \sqrt{\frac{n+1}{\pi}} \lim_{r \rightarrow 1^-} \sum_k a_k \frac{\pi r^{2n+2}}{n+1} \delta_{kn} = \sqrt{\frac{\pi}{n+1}} a_n. \quad (27.1152)$$

Nous effectuons le même genre de calculs pour évaluer  $\|f\|_2^2$  :

$$\|f\|_2^2 = \int_D |f(z)|^2 dz \quad (27.1153a)$$

$$= \lim_{r \rightarrow 1^-} \int_{|z| < r} f(z) \sum_{k=0}^{\infty} \bar{a}_k \bar{z}^k dz \quad (27.1153b)$$

$$= \lim_{r \rightarrow 1^-} \sum_{k=0}^{\infty} \bar{a}_k \int_{|z| < r} f(z) \bar{z}^k dz \quad \text{permuter } \sum \text{ et } \int \quad (27.1153c)$$

$$= \lim_{r \rightarrow 1^-} \sum_{k=0}^{\infty} \bar{a}_k a_k \frac{\pi r^{2k+2}}{k+1} \quad \text{intégrale déjà faite.} \quad (27.1153d)$$

Mais nous savons déjà que  $c_n(f) = \sqrt{\pi/(n+1)}$ , donc ce qui est dans la somme est  $\pi \bar{a}_k a_k / (n+1) = |c_k(f)|^2$ . Nous avons donc

$$\|f\|_2^2 = \lim_{r \rightarrow 1^-} \sum_{k=0}^{\infty} |c_k(f)|^2 r^{2k+2}. \quad (27.1154)$$

La fonction (de  $r$ ) constante  $|c_k(f)|^2$  domine  $|c_k(f)r^{2k+2}|$  tout en ayant une somme (sur  $k$ ) qui converge ; en effet la proposition 25.25 nous indique que  $\sum_k |c_k(f)|^2 \leq \|f\|_2^2$ . Le théorème de la convergence dominée nous permet d'inverser la limite et la somme pour obtenir le résultat attendu :

$$\|f\|_2^2 = \sum_{k=0}^{\infty} |c_k(f)|^2. \quad (27.1155)$$

□