Le Frido 2019,
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 2020. Téléchargeable sur

   `https://laurent.claessens-donadello.eu/pdf/lefrido.pdf`

2. La version la plus complète, contenant des exercices ainsi que de la mathématique de niveau recherche sur

   `https://laurent.claessens-donadello.eu/pdf/giulietta.pdf`

3. Et bien entendu les sources LaTeX

   `https://github.com/LaurentClaessens/mazhe`

Copyright 2011-2019 Laurent Claessens, Carlotta Donadello, Lilian Besson, Bertrand Desmons, and many other 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 David Revoy pour les illustrations de couverture CC-BY,
`https://www.peppercarrot.com/`

ISBN : 979-10-97085-20-9

Chapitre 19

Trigonométrie, isométries

19.1 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 11.269 qui dira que les isométries de cet espace sont des applications linéaires.

19.1.1 Structure du groupe $\mathrm{Isom}(\mathbb{R}^n)$

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 19.2.

Théorème 19.1.
Un peu de structure sur $\mathrm{Isom}(\mathbb{R}^n)$.

(1) *L'application*

$$\begin{aligned}\psi\colon T(n) \times \mathrm{O}(n) &\to \mathrm{Isom}(\mathbb{R}^n)\\ (v, \Lambda) &\mapsto \tau_v \circ \Lambda\end{aligned} \tag{19.1}$$

est une bijection. Ici, $T(n)$ *est le groupe des translations de* $\mathbb{R}^n$.

(2) *Un couple* $(v, \Lambda) \in T(n) \times \mathrm{SO}(n)$ *agit sur* $x \in \mathbb{R}^n$ *par*

$$(v, \Lambda)x = \Lambda x + v \tag{19.2}$$

au sens où $\psi(v, \Lambda)x = \Lambda x + v$.

(3) *En tant que groupes,*

$$\mathrm{Isom}(\mathbb{R}^n) \simeq T(n) \times_\rho \mathrm{O}(n) \tag{19.3}$$

où ρ *représente l'action adjointe de* $\mathrm{O}(n)$ *sur* $T(n)$ *et* $\times_\rho$ *dénotes le produit semi-direct de la définition 2.76.*

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.

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.

Surjection Une isométrie $g \in \mathrm{Isom}(\mathbb{R}^n)$ est une application $g\colon \mathbb{R}^n \to \mathbb{R}^n$ telle que $d(x, y) = d\big(g(x), g(y)\big)$. Dans le cas de $\mathbb{R}^n$ cela se traduit par

$$\|x - y\| = \|g(x) - g(y)\|, \tag{19.4}$$

Vu que $x \mapsto \|x\|$ est une forme quadratique, elle tombe sous le coup du théorème 11.269, 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.

(2) C'est seulement le fait que $(\tau_v \circ \Lambda)x = \tau_v\big(\Lambda x\big) = \Lambda(x) + v$.

(3) Nous allons étudier l'application

$$\psi\colon T(n) \times_\rho O(n) \to \mathrm{Isom}(\mathbb{R}^n). \tag{19.5}$$

Le produit semi-direct est bien définit Il faut montrer que

$$\begin{aligned} \rho\colon O(n) &\to \mathrm{Aut}\big(T(n)\big) \\ \Lambda &\mapsto \mathbf{Ad}(\Lambda) \end{aligned} \tag{19.6}$$

est correcte.

D'abord pour $\Lambda \in O(n)$, nous avons bien $\rho_\Lambda(\tau_v) \in T(n)$ parce qu'en appliquant à $x \in \mathbb{R}^n$,

$$(\Lambda\tau_v\Lambda^{-1})(x) = \Lambda\big(\tau_v(\Lambda^{-1}x)\big) = \Lambda\big(\Lambda^{-1}x + v\big) = x + \Lambda(v) = \tau_{\Lambda(v)}(x). \tag{19.7}$$

Donc $\rho_\Lambda(\tau_v) = \tau_{\Lambda(v)}$.

De plus, $\rho_\Lambda \in \mathrm{Aut}\big(T(n)\big)$ parce que

$$\rho_\Lambda\big(\tau_v \circ \tau_w\big) = \rho_\Lambda(\tau_v) \circ \rho_\Lambda(\tau_v), \tag{19.8}$$

comme on peut aisément vérifier que les deux membres sont égaux à $\tau_{\Lambda(v+w)}$.

ψ est une bijection Cela est déjà vérifié.

ψ est un homomorphisme Nous avons d'une part

$$\psi\big((v,g)(w,h)\big) = \psi\big(v\rho_g(w), gh\big) = \tau_v \circ g \circ \tau_w \circ g^{-1} \circ g \circ h = \tau_v \circ g \circ \tau_w \circ h. \tag{19.9}$$

Et d'autre part,

$$\psi(v,g) \circ \psi(w,h) = \tau_v \circ g \circ \tau_w \circ h, \tag{19.10}$$

ce qui est la même chose.

□

Remarque 19.2.
Notons au passage la loi de groupe sur les couples qui est donnée, pour tout $v, v' \in \mathbb{R}^n$, $\Lambda, \Lambda' \in \mathrm{SO}(n)$, par

$$(v,\Lambda) \cdot (v',\Lambda') = (\Lambda v' + v, \Lambda\Lambda') \tag{19.11}$$

comme le montre le calcul suivant :

$$\begin{aligned} (v,\Lambda) \cdot (v',\Lambda')x &= (v,\Lambda)(\Lambda' x + v') && \text{(19.12a)} \\ &= \Lambda\Lambda' x + \Lambda v' + v && \text{(19.12b)} \\ &= (\Lambda v' + v, \Lambda\Lambda')x. && \text{(19.12c)} \end{aligned}$$

Proposition 19.3 ([313])**.**
Soient $n \geqslant 1$ et R un élément de $\mathrm{O}(n)$ de déterminant -1 tels que $R^2 = \mathrm{Id}$. En posant $C_2 = \{\mathrm{Id}, R\}$ nous avons

$$\mathrm{O}(n) = \mathrm{SO}(n) \times_\rho C_2 \tag{19.13}$$

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

$$\begin{aligned} \psi\colon \mathrm{SO}(n) \times C_2 &\to \mathrm{O}(n) \\ (A,h) &\mapsto Ah. \end{aligned} \tag{19.14}$$

est un isomorphisme.

Injectif Soient $A, B \in \mathrm{SO}(n)$ et $h, k \in C_2$ tels que $\psi(A,h) = \psi(B,k)$, c'est-à-dire tels que $Ah = Bk$. Vu que $\det(A) = \det(B) = 1$ 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 $A = B$.

Surjectif Soit $X \in \mathrm{O}(n)$. Si $\det(X) = 1$ alors $X \in \mathrm{SO}(n)$ et $X = \psi(X, \mathbb{1})$. Si par contre $\det(X) = -1$ alors $XR \in \mathrm{SO}(n)$ parce que $\det(XR) = 1$ et nous avons

$$\psi(XR, R) = XR^2 = X. \tag{19.15}$$

Homomorphisme Nous avons

$$\psi\Big((A,h)(B,k)\Big) = \psi\big(A\rho_h(B), hk\big) = A(hBh^{-1})hk = AhBk, \tag{19.16}$$

tandis que

$$\psi(A,h)\psi(B,k) = AhBk, \tag{19.17}$$

qui est la même chose.

□

Lemme 19.4 ([314]).
Si $n \geqslant 3$, alors toute droite est intersection de deux plans non isotropes.

Proposition 19.5 ([313]).
Si une isométrie de $\mathbb{R}^n$ fixe un ensemble F de points, alors elle fixe l'espace affine engendrée par F.

Démonstration. Soit $f \in \mathrm{Isom}(\mathbb{R}^n)$ fixant F. Par le théorème 11.269, c'est une application affine et l'ensemble $\mathrm{Fix}(f)$ des points fixés par f est un sous-espace affine de $\mathbb{R}^n$, grâce à la proposition 10.55.

Donc $\mathrm{Fix}(f)$ est un espace affine contenant F. Vu que l'espace affine engendré par F est l'intersection de tous les espaces affines contenant F, il est en particulier contenu dans $\mathrm{Fix}(f)$. □

Corollaire 19.6.
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 19.5, l'espace affine engendré par F. Or tout point fixé par h est un point sur lequel g et f coïncident. □

19.1.2 Classification des isométries de $\mathbb{R}$

Définition 19.7.
*Soit $x \in \mathbb{R}$; nous notons σ_x la **réflexion** par rapport à x, c'est-à-dire*

$$\sigma_x(y) = 2x - y. \tag{19.18}$$

Théorème 19.8 ([313]).
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 réflexions),*
- *les réflexions,*
- *les translations (2 réflexions)*

Démonstration. Nous divisions la preuve en fonction du nombre de points fixés par l'isométrie $f \in \mathrm{Isom}(\mathbb{R})$.

f fixe deux points distincts Alors elle fixe l'espace affine engendrée par ces deux points par la proposition 19.5. Donc f fixe tout $\mathbb{R}$ et est l'identité.

f fixe un unique point Soit x l'unique point fixé par f et considérons $y \neq x$. Vu que $x = f(x)$ et que f est une isométrie,

$$d\big(x, f(y)\big) = d\big(f(x), f(y)\big) = d(x, y). \tag{19.19}$$

Donc $f(y)$ est à égale distance de x que 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 $f = \sigma_x$.

f n'a pas de points fixes 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

$$g(x) = \sigma_y\big(f(x)\big) = 2y - f(x) = x. \tag{19.20}$$

Donc soit g est l'identité soit g est une réflexion (par les points précédents). La possibilité $g = \mathrm{Id}$ est exclue parce que cela ferait $f = \sigma_y$ alors que f n'a pas de points fixes. Donc g est une réflexion ; et comme x est un point fixe de g nous avons $g = \sigma_x$. Au final

$$f = \sigma_y \circ \sigma_x. \tag{19.21}$$

Montrons que cela implique que f est une translation :

$$\sigma_y\sigma_x(z) = \sigma_y(2x - z) = 2y - 2x + z = z + 2(y - x). \tag{19.22}$$

Donc $\sigma_y \circ \sigma_x$ est la translation de vecteur $2(y - x)$.

□

19.1.3 Segment, plan médiateur et équidistance

Lemme 19.9.
Un point M est sur la médiatrice du segment $[A, B]$ si et seulement si $\|M - A\| = \|M - B\|$.

Lemme 19.10.
Soient A et B de points de $\mathbb{R}^3$. Alors le plan médiateur du segment $[A, B]$ est le lieu des points de $\mathbb{R}^3$ à être équidistants de A et B.

Démonstration. Nous nommons σ ce plan.

Soit X un point équidistant de A et B. Alors dans le plan (A, B, X), le triangle ABX est isocèle en X, et la hauteur issue de X coupe perpendiculairement $[A, B]$ en son milieu. Cela prouve que X est dans le plan médiateur du segment $[A, B]$ (lemme 19.9).

Mettons au contraire que X est dans le plan médiateur de $[A, B]$. Nous avons $(X, M) \perp (A, B)$. Donc le triangle A, B, X est isocèle en X et donc X est équidistant de A et B. □

19.1.4 Isométries du tétraèdre régulier

Définition 19.11.
*Un polyèdre **régulier** est un polyèdre dont les faces sont des polygones réguliers identiques dont tous les sommets joignent le même nombre d'arrêtes.*

Le **tétraèdre** est une pyramide à base triangulaire dont toutes les faces sont des triangles équilatéraux.

Proposition 19.12 (Isométries affines du tétraèdre régulier)**.**
Soient T un tétraèdre régulier et $\mathrm{Iso}(T)$ son groupe d'isométries affines (définition 10.58). Alors

$$\mathrm{Iso}(T) \simeq S_4 \tag{19.23}$$

où S_4 est le groupe des permutations de quatre objets.

Démonstration. Commençons par prouver qu'une isométrie préserve les sommets : l'image d'un sommet est un sommet. Pour cela nous considérons $g \in \mathrm{Iso}(T)$ et nous supposons que l'image d'un sommet x soit à l'intérieur d'une arrête. Soient $g(y)$ et $g(z)$ deux points distincts de cette arrête situés à égale distance de $g(x)$. Cela est possible parce que g est une bijection de $\mathbb{R}^3$. Aussi : $y \neq z$. Mais une application affine préserve l'alignement (vous ne le croyez pas ? regardez la forme donnée par le lemme (10.53)), donc x, y et z foment un triangle isocèle en x de points alignés et appartenant à T. Cela est impossible si x est un sommet.

Donc l'image d'un sommet est un sommet. Si nous numérotons les sommets $x_1,\dots, x_4$, nous obtenons un morphisme de groupe $\varphi\colon \mathrm{Isom}(T) \to S_4$ 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.

Le morphisme φ 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 $\mathbb{R}^4 \to \mathbb{R}^3$ fixant 4 point est l'identité par le lemme 10.50. Donc $g_1^{-1}g_2 = \mathrm{Id}$, ce qui prouve que $g_1 = g_2$. Vous noterez que nous utilisons l'unicité de l'inverse dans un groupe.

φ est surjectf Nous savons que S_4 est engendré par les transpositions (proposition 2.70). Or les transpositions sont dans l'image de φ. En effet, notons les sommets de notre tétraèdre par A, B, D et D et considérons la transposition $A \leftrightarrow B$. Elle est l'image par φ de la réflexion selon le plan σ, médiateur du segment $[A, B]$. Pour nous assurer de cela, nous devons nous assurer que C et D appartiennent à σ. Cela est le contenu du lemme 19.10.

Conclusion L'application φ est un morphisme bijectif, c'est-à-dire un isomorphisme.

□

19.1.5 Représentation de S_4 via les isométries du tétraèdre

19.13.
Lorsque le tétraèdre a son barycentre en l'origine de $\mathbb{R}^3$, l'isomorphisme $\varphi\colon \mathrm{Iso}(T) \to S_4$ donne une représentation de dimension 3 de S_4. Nous avons calculé les caractères de S_4 en la section 17.5 sans avoir besoin de savoir que l'une des représentations de dimension 3 est cella que nous venons de trouver via le groupe des isométries du tétraèdre. Nous allons cependant également y calculer les caractères de la représentations φ, pour le sport.

Une des représentations trouvées (la représentation ρ_s) peut être vue comme le groupe $\mathrm{Iso}(T)$ des isométries affine du tétraèdre grâce à la proposition 19.12 qui donne un isomorphisme de groupe $S_4 \simeq \mathrm{Iso}(T)$ 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 $\mathrm{Iso}(T)$ sont des applications linéaires parce que

— les affinités laissent invariantes les barycentres (proposition 10.40),

— les affinités qui laissent l'origine invariante sont linéaires (corollaire 10.54).

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 2.68).

Pour tout cela nous allons considérer un tétraèdre dont le centre (isobary) est en $(0, 0, 0)$ et une base de $\mathbb{R}^3$ formée de trois sommets e_1, e_2 et e_3. Vu que l'isobarycentre des quatre sommes est en $(0, 0, 0)$, le quatrième somme 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 le 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 e_4 sont équidistants de e_1 et e_2. Le lemme 19.10 dit qu'alors ces points dont partie du plan médiateur.

Dans notre base, la matrice de la transposition précédemment nommée (12) est

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{19.24}$$

dont la trace est 1. Donc $\chi_s(12) = 1$.

Les bitranspositions La bitransposition (12)(34) est le produit des transpositions selon les plans médiateur 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 :

$$\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} \tag{19.25}$$

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

$$\chi_s\big((12)(34)\big) = -1. \tag{19.26}$$

Les 3-cycles La symétrie qui permute cycliquement les points e_1, e_2 et e_3 est la rotation d'angle $2\pi/3$ 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 $2\pi/3$ dans n'importe quelle base. Par exemple :

$$\chi_s\big((12)(34)\big) = \operatorname{Tr}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2\pi/3) & \sin(2\pi/3) \\ 0 & -\sin(2\pi/3) & \cos(2\pi/3) \end{pmatrix} = 1 + 2\cos(2\pi/3) = 0. \tag{19.27}$$

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 :

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \tag{19.28}$$

dont la trace est manifestement nulle : $\chi_s\big((123)\big) = 0$.

Le 4-cycle Il fait $e_1 \mapsto e_2 \mapsto e_3 \mapsto e_4 \mapsto e_1$, dont la matrice est

$$\begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \tag{19.29}$$

et la trace est $\chi_s\big((1,2,3,4)\big) = -1$.

Nous avons retrouvé les caractères de la représentation ρ_s, et nous pouvons vérifier qu'elle est irréductible.

19.2 Transformations de Lorentz

Nous considérons dans cette section un nombre réel $c > 0$ ainsi l'espace $\mathbb{R}^2$ muni du produit pseudo-scalaire [1] donné par la matrice

$$\eta = \begin{pmatrix} c^2 & 0 \\ 0 & -1 \end{pmatrix}. \tag{19.30}$$

Et pour faire plus vrai, nous notons (x_0, x_1) les coordonnées sur $\mathbb{R}^2$. Ainsi

$$x \cdot y = c^2 x_0 y_0 - x_1 y_1. \tag{19.31}$$

Nous insistons sur le fait que cela n'est pas un produit scalaire.

1. Définition 11.6.

Lemme 19.14 ([1]).
Soit $c > 0$. L'application

$$\begin{aligned}\varphi\colon]-c,c[&\to \mathbb{R}\\ v &\mapsto \frac{-v/c}{\sqrt{1-\frac{v^2}{c^2}}}\end{aligned} \tag{19.32}$$

est une bijection.

Démonstration. Nous commençons par mentionner le fait que φ est continue du fait que le dénominateur ne s'annule pas. Une petite étude fonction montre que

$$\lim_{v\to -c}\varphi(v) = \infty, \tag{19.33}$$

et

$$\lim_{v\to c}\varphi(v) = -\infty, \tag{19.34}$$

et

$$\varphi'(v) = -\frac{1}{c\sqrt{1-\frac{v^2}{c^2}}} - \frac{v^2/c^3}{\left(1-\frac{v^2}{c^2}\right)^{3/2}} < 0. \tag{19.35}$$

Tout cela fait que φ est bijective (entre autres par le théorème des valeurs intermédiaires 13.50 et la théorème dérivée et croissance 13.124). □

Théorème 19.15.
Soit une bijection[2] *$f\colon \mathbb{R}^2 \to \mathbb{R}^2$ telle que*

$$f(x)\cdot f(y) = x\cdot y \tag{19.36}$$

pour tout $x, y \in \mathbb{R}^2$. Alors :

(1) f est linéaire.

(2) Il existe un unique choix de $(x, \sigma_1, \sigma_2) \in \mathbb{R}\times\{\pm 1\}\times\{\pm 1\}$ tel que la matrice de f ait la forme

$$f = \begin{pmatrix} \sigma_1\cosh(\xi) & \frac{\sigma_1\sigma_2}{c}\sinh(\xi)\\ c\sinh(\xi) & \sigma_2\cosh(\xi)\end{pmatrix}. \tag{19.37}$$

(3) Il existe un unique $v \in]-c,c[$ tel que la matrice de f ait la forme

$$f = \begin{pmatrix} \frac{\sigma_1}{\sqrt{1-\frac{v^2}{c^2}}} & -\frac{\sigma_1\sigma_2}{c^2}\frac{v}{\sqrt{1-\frac{v^2}{c^2}}}\\ \frac{-v}{\sqrt{1-\frac{v^2}{c^2}}} & \frac{\sigma_2}{\sqrt{1-\frac{v^2}{c^2}}}\end{pmatrix}. \tag{19.38}$$

Démonstration. Le fait que f doive être linéaire est la proposition 11.269. Nous posons

$$A = \begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix} \tag{19.39}$$

et, conformément à la proposition 11.273 nous imposons $A^t\eta A = \eta$. Après un petit produit matriciel nous obtenons :

$$\begin{pmatrix} c^2\alpha^2-\gamma^2 & c^2\alpha\beta-\gamma\delta\\ c^2\alpha\beta-\gamma\delta & c^2\beta^2-\delta^2\end{pmatrix} = \begin{pmatrix} c^2 & 0\\ 0 & -1\end{pmatrix}. \tag{19.40}$$

Voila 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 :

$$\begin{cases} c^2\alpha^2-\gamma^2 = c^2 & \text{(19.41a)}\\ c^2\alpha\beta-\gamma\delta = 0 & \text{(19.41b)}\\ c^2\beta^2-\delta^2 = 1. & \text{(19.41c)}\end{cases}$$

2. À mon avis, il y a moyen d'affaiblir cette hypothèse. Écrivez-moi si vous avez une idée.

C'est le moment d'utiliser la proposition 16.91. La relation (19.41a) donne

$$\alpha^2 - \left(\frac{\gamma}{c}\right)^2 = 1, \tag{19.42}$$

ce qui implique l'existence (unique) de $\xi_1 \in \mathbb{R}$ et $\sigma_1 \in \{\pm 1\}$ tels que

$$\gamma = c\sinh(\xi_1) \tag{19.43a}$$
$$\alpha = \sigma_1 \cosh(\xi_1). \tag{19.43b}$$

La relation (19.41c) implique quant à elle l'existence de $\xi_2 \in \mathbb{R}$ et $\sigma_2 \in \{\pm 1\}$ tels que

$$\delta = \sigma_2 \cosh(\xi_2) \tag{19.44a}$$
$$\beta = \frac{1}{c}\sinh(\xi_2). \tag{19.44b}$$

Nous substituons maintenant toutes les valeurs (19.43) et (19.44) dans (19.41b). Cela donne

$$\sigma_1 \cosh(\xi_1)\sinh(\xi_2) = \sinh(\xi_1)\cosh(\xi_2). \tag{19.45}$$

Nous mettons cette relation au carré et nous substituons $\cosh(\xi_1)^2 = 1 + \sinh^2(\xi_1)$. Ce que nous trouvons est

$$\sinh(\xi_1)^2 = \sinh(\xi_2)^2, \tag{19.46}$$

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 ξ au lieu de ξ_1.

Nous remettons la valeur $\xi = \xi_1 = \sigma_3\xi_2$ dans l'équation (19.45) en tenant compte du fait que sinh est impaire et cosh est paire :

$$\sigma_1\sigma_3 \cosh(\xi)\sinh(\xi) = \sigma_2 \sinh(\xi)\cosh(\xi). \tag{19.47}$$

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

$$A = \begin{pmatrix} \sigma_1 \cosh(\xi) & \frac{\sigma_1\sigma_2}{c}\sinh(\xi) \\ c\sinh(\xi) & \sigma_2\cosh(\xi) \end{pmatrix}. \tag{19.48}$$

Nous utilisons à présent la bijection du lemme 19.14. Il existe un unique $v \in]-v, v[$ tel que $\sinh(\xi) = \varphi(v)$. En utilisant $\cosh(\xi)^2 = 1 + \varphi(v)^2$, nous trouvons

$$\cosh(\xi)^2 = \frac{1}{1 - \frac{v^2}{c^2}}. \tag{19.49}$$

Mais comme le cosinus hyperbolique est toujours strictement positif, nous pouvons prendre la racine carré des deux côtés :

$$\cosh(\xi) = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{19.50}$$

En substituant dans (19.48), nous trouvons le résultat annoncé. □

19.3 Trigonométrie

19.3.1 Définitions, périodicité et quelques valeurs remarquables

Proposition-définition 19.16 (Défintion du cosinus et du sinus)**.**
La série

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}x^{2n} \tag{19.51}$$

définit une fonction $\cos\colon \mathbb{R} \to \mathbb{R}$ *de classe* C^∞. *Nous l'appelons* ***cosinus***.

La série

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \tag{19.52}$$

définit une fonction $\sin\colon \mathbb{R} \to \mathbb{R}$ *de classe* C^∞. *Nous l'appelons* ***sinus***.

Démonstration. La série entière définissant $\cos(x)$ a pour coefficients

$$a_n = \begin{cases} 0 & \text{si } n \text{ est impair} \\ \frac{(-1)^{n/2}}{n!} & \text{si } n \text{ est pair.} \end{cases} \tag{19.53}$$

Nous pouvons la majorer par la série entière donnée par les coefficients

$$b_n = \begin{cases} 1/n! & \text{si } n \text{ est impair} \\ \frac{(-1)^{n/2}}{n!} & \text{si } n \text{ est pair.} \end{cases} \tag{19.54}$$

Quelle que soit la parité de k nous avons toujours

$$\left|\frac{b_{k+1}}{b_k}\right| = \frac{1}{k+1}, \tag{19.55}$$

de telle sorte que la formule d'Hadamard (16.31) nous donne $R = \infty$ pour la série $\sum_{k=0}^{\infty} b_k x^k$. A fortiori[3] 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^∞ sur $\mathbb{R}$, il faut invoquer le corollaire 16.27. □

Par substitution directe dans les séries, nous avons immédiatement

$$\cos(0) = 1 \tag{19.56a}$$
$$\sin(0) = 1. \tag{19.56b}$$

Lemme 19.17.
En ce qui concerne la dérivation, nous avons

$$\sin' = \cos \tag{19.57a}$$
$$\cos' = -\sin. \tag{19.57b}$$

Démonstration. Il s'agit de se permettre de dériver terme à terme (proposition 16.25) les séries qui définissent le sinus et le cosinus. □

Lemme 19.18.
Les fonctions sinus et cosinus vérifient

$$\cos^2(x) + \sin^2(x) = 1 \tag{19.58}$$

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

Démonstration. Posons $f(x) = \sin^2(x) + \cos^2(x)$ et dérivons :

$$f'(x) = 2\sin(x)\cos(x) + 2\cos(x)(-)\sin(x) = 0. \tag{19.59}$$

La fonction f est donc constante par le corollaire 13.133. Nous avons donc pour tout x :

$$f(x) = f(0) = \sin^2(0) + \cos^2(0) = 1. \tag{19.60}$$

Le dernier calcul s'obtient en substituant directement x par zéro dans les séries : $\sin(0) = 0$ et $\cos(0) = 1$. □

3. Remarque 16.9.

19.3.2 Fonction puissance (pour les complexes)

La fonction puissance a déjà fait l'objet de nombreuses définitions et extensions. Voir le thème 32. Nous allons maintenant définir a^z pour $a > 0$ et $z \in \mathbb{C}$.

Définition 19.19.
Pour le nombre $e \in \mathbb{R}$ et le nombre imaginaire pur iy ($y \in \mathbb{R}$), nous définisson

$$e^{iy} = \exp(iy) \tag{19.61}$$

où exp *est la série usuelle de la définition 16.37. Pour un nombre complexe général $x + yi$ nous définissons*

$$e^{x+iy} = e^x e^{iy}. \tag{19.62}$$

Et enfin, si $a > 0$ et si $z \in \mathbb{C}$ nous définissons

$$a^z = e^{z \ln(a)}, \tag{19.63}$$

la fonction logarithme ici étant celle $\ln\colon]0, \infty[\to \mathbb{R}$ *définie par la proposition 16.58.*

Si $z \in \mathbb{C}$ et si $n \in \mathbb{Z}$, la définition de z^n ne pose pas de problèmes, c'est la définition 1.59.

19.20.
Soit $z = x + iy \in \mathbb{C}$. L'exponentielle $\exp(x+yi)$ est déjà définie en 16.37 ; elle est la fonction donnée par

$$\begin{aligned} \exp\colon \mathbb{C} &\to \mathbb{C} \\ z &\mapsto \sum_{n=0}^{\infty} \frac{z^n}{n!}. \end{aligned} \tag{19.64}$$

Proposition 19.21.
Le rayon de convergence[4] *de la série exponentielle est infini.*

Démonstration. La formule de Hadamard de la proposition 13.340 est à utiliser avec $a_k = j!$. Nous avons

$$\frac{1}{R} = \lim_{k\to\infty} \left| \frac{(n+1)!}{n!} \right| = \lim_{k\to\infty} (n+1) = \infty. \tag{19.65}$$

Donc $R = \infty$. □

Proposition 19.22.
Pour tout $z \in \mathbb{C}$ nous avons

$$\exp(z) = e^z. \tag{19.66}$$

Proposition 19.23 ([315])**.**
Quelques propriétés de l'exponentielle.

(1) *Le 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) $(\exp(z))^n = \exp(nz)$.

Démonstration. La proposition 19.21 nous enseigne que le rayon de convergence est infini. La fonction ainsi définie est alors continue par la proposition 13.291.

4. Définition 16.8.

Les séries $\exp(z)$ et $\exp(w)$ ayant un rayon de convergence infini nous pouvons utiliser le produit de Cauchy (théorème 16.15) :

$$e^z e^w = \sum_{n=0}^{\infty} \left(\sum_{i+j=n} \frac{z^i w^j}{i!j!} \right) \tag{19.67a}$$

$$= \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} \frac{z^i w^{n-i}}{i!(n-i)!} \right) \tag{19.67b}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} z^i w^{n-i} \tag{19.67c}$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} (z+w)^n \tag{19.67d}$$

$$= \exp(z+w). \tag{19.67e}$$

Nous avons utilisé la formule du binôme (proposition 3.40).

Les autres propriétés énoncées sont des corollaires :

$$e^z e^{-z} = e^0 = 1. \tag{19.68}$$

□

D'autres propriétés de l'exponentielle sur $\mathbb{C}$, entre autres l'holomorphie, sont données dans le théorème 27.47.

Lemme 19.24 ([1])**.**
Soient $a > 0$, $z \in \mathbb{C}$ et $n \in \mathbb{Z}$. Alors

$$(a^z)^n = a^{nz}. \tag{19.69}$$

19.3.3 Formules de trigonométrie

Le lemme suivant est un premier pas pour la paramétrisation du cercle dont nous parlerons dans la proposition 19.60.

Lemme 19.25.
Nous avons la formule

$$e^{ix} = \cos(x) + i\sin(x) \tag{19.70}$$

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

En particulier pour tout x, nous avons $|e^{ix}| = 1$.

Démonstration. La définition de l'exponentielle sur $\mathbb{C}$ est la définition 16.37. Cette définition fonctionne parce que $\mathbb{C}$ est une algèbre normée, et que $\mathbb{C}$ est un $\mathbb{C}$-module vérifiant l'inégalité $|zz'| \leqslant |z||z'|$ (en l'occurrence, une égalité).

Nous remarquons que 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,

$$e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} = \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{y^{2n+1}}{(2n+1)!} = \cos(y) + i\sin(y). \tag{19.71}$$

Nous avons utilisé le fait que $i^{2n} = (-1)^n$ et $i^{2n+1} = i(-1)^n$. □

Lemme 19.26.
Nous avons les formules d'addition d'angles[5]

$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \tag{19.72a}$$
$$\sin(a+b) = \cos(a)\sin(b) + \sin(a)\cos(b) \tag{19.72b}$$
$$\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b) \tag{19.72c}$$

pour tout a, b réels.

Démonstration. Nous utilisons la formule d'addition dans l'exponentielle, proposition (16.203) et la formule (19.70) avant de séparer les parties réelles et imaginaires :

$$e^{i(a+b)} = e^{ia}e^{ib} = \cos(a)\cos(b) - \sin(a)\sin(b) + i\big(\cos(a)\sin(b) + \sin(a)\cos(b)\big). \tag{19.73}$$

Cela est également égal à

$$\cos(a+b) + i\sin(a+b). \tag{19.74}$$

En identifiant les parties réelle et imaginaires, nous obtenons les formules (19.72a) et (19.72c) annoncées.

Pour la formule (19.72c), 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 (19.72a) avec $-b$ au lieu de b. □

Corollaire 19.27.
Les formules suivantes pour les duplications d'angles s'ensuivent :

$$\cos(2a) = \cos^2(a) - \sin^2(a) \tag{19.75a}$$
$$\sin(2a) = 2\cos(a)\sin(a). \tag{19.75b}$$

Démonstration. Poser $b = a$ dans les relations du lemme 19.26. □

Lemme 19.28.
Un sous-groupe de $(\mathbb{R}, +)$ est soit dense dans $\mathbb{R}$ soit de la forme $p\mathbb{Z}$ pour un certain réel $p \neq 0$.

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 $\mathbb{R}$). 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 $p \geqslant d$ (parce que $0 \in A$ de toutes façons).

Vu que 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. □

Proposition-définition 19.29 (Périodicité, le nombre π[316])**.**
Plusieurs choses à propos de la périodicité de la fonction cos*.*

(1) La fonction cos *est périodique.*

(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 *:*

$$2\pi = \min\{T > 0 \text{ tel que } \cos(x+T) = \cos(x)\ \forall x\}. \tag{19.76}$$

Démonstration. Plusieurs étapes.

5. Rien ne nous empêche de donner ce nom à ces formules, mais seriez vous capable de définir précisément le mot « angle » ?

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 13.66 :

$$\lim_{x\to\infty} \sin(x) = \ell \tag{19.77}$$

pour un certain $\ell > 0$. Avec ça nous avons aussi (pour cause de dérivée) $\lim_{x\to\infty} \sin'(x) = 0$, c'est-à-dire $\lim_{x\to\infty} \cos(x) = 0$. Mais vu que $\cos^2(x) + \sin^2(x) = 1$ nous en déduisons que $\lim_{x\to\infty} \sin(x) = 1$. Mézalor $\lim_{x\to\infty} \cos'(x) = -1$, ce qui donne que la fonction cos n'est pas bornée. Cela est impossible. Nous en déduisons que $\cos(x)$ n'est pas toujours positive.

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 (19.72),

$$\cos(2r) = \cos^2(r) - \sin^2(r) = -1 \tag{19.78a}$$
$$\sin(2r) = 2\cos(r)\sin(r) = 0. \tag{19.78b}$$

et par conséquent

$$\cos(4r) = \cos^2(2r) - \sin^2(2r) = 1 \tag{19.79a}$$
$$\sin(4r) = 2\cos(2r)\sin(2r) = 0. \tag{19.79b}$$

Donc $T = 4r$ fonctionne.

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

$$\cos(x + T) = \cos(x)\cos(T) - \sin(x)\sin(T) = \cos(x) \tag{19.80}$$

et donc tel que

$$\cos(x)\big(\cos(T) - 1\big) = \sin(x)\sin(T). \tag{19.81}$$

Nous dérivons cette équation :

$$-\sin(x)\big(\cos(T) - 1\big) = \cos(x)\sin(T). \tag{19.82}$$

Nous multiplions chacune des deux équations (19.81) et (19.82) par $\sin(x)$ et $\cos(x)$ pour obtenir les quatre relations suivantes :

$$\cos^2(x)\big(\cos(T) - 1\big) - \sin(x)\cos(x)\sin(T) = 0 \tag{19.83a}$$
$$-\sin(x)\cos(x)\big(\cos(T) - 1\big) - \cos^2(x)\sin(T) = 0 \tag{19.83b}$$
$$\sin(x)\cos(x)\big(\cos(T) - 1\big) - \sin^2(x)\sin(T) = 0 \tag{19.83c}$$
$$\sin^2(x)\big(\cos(T) - 1\big) - \sin(x)\cos(x)\sin(T) = 0 \tag{19.83d}$$

En faisant (19.83a) moins (19.83d) nous trouvons $\cos(T) = 1$. Et en sommant (19.83b) avec (19.83c) nous avons $-\sin(T) = 0$.

Si $T > 0$ est tel que $\sin(T) = 0$ et $\cos(T) = 1$ Alors les formules d'addition d'angle du lemme 19.26 donnent tout de suite

$$\cos(x + T) = \cos(x). \tag{19.84}$$

À de niveau nous croyons avoir prouvé que cos était périodique et que la période est donnée par

$$\min\{T > 0 \text{ tel que } \sin(T) = 0, \cos(T) = 1\}. \tag{19.85}$$

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 19.28 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, soit $t \in \mathbb{R}$ et une suite (t_n) dans P telle que $t_n \to t$. Pour tout x et tout n nous avons

$$\cos(x + t_n) = \cos(x), \tag{19.86}$$

Vu que 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. □

Proposition 19.30.
La fonction sin *est périodique de période* π *et*

$$2\pi = \min\{T > 0 \text{ tel que } \sin(T) = 0, \cos(T) = 1\}. \tag{19.87}$$

19.31.
Notons que tout ceci ne nous donne pas la plus petite indication d'ordre de grandeur de la valeur de π. Cela peut encore être 0.1 autant que 500.

Proposition 19.32 ([316])**.**
Des propriétés à la chaîne à propos des sinus, cosinus et de leurs périodes.

(1) *Le nombre* 2π *est le plus petit nombre strictement positif tel que*

$$\begin{cases} \cos(2\pi) = 1 & \text{(19.88a)} \\ \sin(2\pi) = 0. & \text{(19.88b)} \end{cases}$$

(2) *Les fonctions* sin *et* cos *sont périodiques de période* 2π*.*

(3) *Nous avons* $\cos(\pi) = -1$ *et* $\sin(\pi) = 0$*.*

(4) *Pour tout* $a \in \mathbb{R}$ *nous avons*

$$\begin{aligned} \cos(a + \pi) &= -\cos(\pi) & \text{(19.89a)} \\ \sin(a + \pi) &= -\sin(\pi). & \text{(19.89b)} \end{aligned}$$

(5) *Le nombre* π *est le plus petit* $T > 0$ *tel que* $\cos(T) = -1$ *et* $\sin(T) = 0$*.*

(6) *Nous avons*

$$\begin{cases} \cos(\pi/2) = 0 & \text{(19.90a)} \\ \sin(\pi/2) = 1. & \text{(19.90b)} \end{cases}$$

(7) *Nous avons les formules*

$$\begin{cases} \cos(x + \pi/2) = -\sin(x) & \text{(19.91a)} \\ \sin(x + \pi/2) = \cos(x) & \text{(19.91b)} \end{cases}$$

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

(8) *Le nombre* $\pi/2$ *est le plus petit* T *vérifiant* $\sin(T) = 1$, $\cos(T) = 0$*.*

(9) *Nous avons les valeurs*

$$\begin{cases} \cos(\frac{3\pi}{2}) = 0 & \text{(19.92a)} \\ \sin(\frac{3\pi}{2}) = -1. & \text{(19.92b)} \end{cases}$$

(10) *Le nombre* $\pi/2$ *est le plus petit* $T > 0$ *tel que* $\cos(T) = 0$*.*

Démonstration. C'est parti.

(1) Le fond de la proposition 19.29 est que toutes les périodes $T > 0$ vérifient $\cos(T) = 1$ et $\sin(T) = 0$. La définition de π 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

$$\begin{aligned}\sin(x+T) &= -\cos'(x+T) && (19.93a)\\ &= -\lim_{\epsilon\to 0}\frac{\cos(x+T+\epsilon)-\cos(x+T)}{\epsilon} && (19.93b)\\ &= -\lim_{\epsilon\to 0}\frac{\cos(x+\epsilon)-\cos(x)}{\epsilon} && (19.93c)\\ &= -\cos'(x) && (19.93d)\\ &= \sin(x). && (19.93e)\end{aligned}$$

Nous déduisons que toute période de cos est une période de sin. De la même façon, nous pouvons prouver le contraire : toute période de sin est une période de cos.

(3) D'un côté nous avons

$$\cos(2\pi) = \cos^2(\pi) - \sin^2(\pi) = 1 \tag{19.94}$$

parce que $\cos(2\pi) = \cos(0) = 1$. Vu que $\cos(\pi)$ 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 2π est le plus petit T vérifiant $\cos(T) = 1$, $\sin(T) = 0$. Donc avoir $\cos(\pi) = 1$ n'est pas possible. Nous concluons

$$\begin{cases}\cos(\pi) = -1 & (19.95a)\\ \sin(\pi) = 0. & (19.95b)\end{cases}$$

(4) Il s'agit d'utiliser les formules d'addition d'angles du lemme 19.26 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

$$\begin{aligned}\cos(a+\pi) &= -\cos(\pi) = 1 && (19.96a)\\ \sin(a+\pi) &= -\sin(\pi) = 0, && (19.96b)\end{aligned}$$

ce qui donnerait $a + \pi \in]\pi, 2\pi[$ dont le cosinus est 1 et le sinus est zéro. Mais nous savons déjà que 2π est le minimum pour cette propriété.

(6) Nous avons

$$-1 = \cos(\pi) = \cos^2(\pi/2) - \sin^2(\pi/2), \tag{19.97}$$

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 $\sin'(0) = \cos(0) = 1$, donc il existe $\epsilon > \epsilon$ tel que pour tout $x \in]0, \epsilon[$ nous avons $0 < \cos(x) < 1$ et $0 \sin(x) < 1$. Nous choisissons ϵ plus petit que $\pi/2$.

Supposons que $\sin(\pi/2) = -1$. Le théorème des valeurs intermédiaires 13.50 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 vu que 2π est minium pour avoir cos = 1 et sin = 0 nous devons avoir $\cos(x_0) = -1$. Alors nous avons aussi

$$\begin{aligned}\cos(x_0+\pi) &= \cos(x_0)\cos(\pi) - \sin(x_0)\sin(\pi) = -\cos(x_0) = 1 && (19.98a)\\ \sin(x_0+\pi) &= \cos(x_0)\sin(\pi) + \sin(x_0)\cos(\pi) = \sin(x_0) = 0. && (19.98b)\end{aligned}$$

Encore une fois par minimalité de 2π, 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 du fait que $\cos(\pi/2) = 0$ 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 (19.91) nous avons

$$\cos(x_0 + \pi/2) = -1 \tag{19.99a}$$
$$\sin(x_0 + \pi/2) = 0, \tag{19.99b}$$

avec $x_0 + \pi/2 < \pi$. Cela contredirait la minimalité de π.

(9) Il s'agit d'utiliser les formules (19.91) :

$$\cos(\frac{3\pi}{2}) = \cos(\pi + \pi/2) = -\sin(\pi) = 0 \tag{19.100a}$$
$$\sin(\frac{3\pi}{2}) = \sin(\pi + \pi/2) = \cos(\pi) = -1. \tag{19.100b}$$

(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.

□

Corollaire 19.33.
Des nombres $x, y \in \mathbb{R}$ vérifient $e^{ix} = e^{iy}$ si et seulement si il existe $k \in \mathbb{Z}$ tel que $y = x + 2k\pi$.

Tout cela nous permet de calculer quelque valeurs remarquables de cosinus et sinus ainsi que d'écrire le tableau de variations de sinus et cosinus.

Lemme 19.34.
Nous avons les valeurs remarquables

$$\sin(\frac{\pi}{4}) = \cos(\frac{\pi}{4}) = \frac{\sqrt{2}}{2}. \tag{19.101}$$

Démonstration. La relation (19.75b) donne

$$0 = \cos(\pi/2) = \cos^2(\pi/4) - \sin^2(\pi/4). \tag{19.102}$$

Donc $\cos^2(\pi/4) = \sin^2(\pi/4)$. Mais vu que $\sin(\pi/4)$ et $\cos(\pi/4)$ sont positifs, ils sont égaux.

Mais $\sin^2(\pi/4) + \cos^2(\pi/4) = 1$. Donc le nombre $x = \cos(\pi/4) = \sin(\pi/4)$ vérifie l'équation $2x^2 = 1$, donc l'unique solution positive est $x = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$. □

Lemme 19.35.
Nous avons la valeur remarquable

$$\cos(\frac{\pi}{3}) = \frac{1}{2}. \tag{19.103}$$

Démonstration. Il faut utiliser la formule (19.72a) 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

$$4x^3 - 3x + 1 = 0. \tag{19.104}$$

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

$$4x^3 - 3x + 1 = (2x - 1)^2(x + 1). \tag{19.105}$$

Donc $1/2$ est de multiplicité 2 et -1 de multiplicité 1. Le théorème 3.177 nous dit qu'il n'y a alors pas d'autres racines que ces deux là [6].

Nous en déduisons que la valeur de $\cos(\pi/3)$ est soit $1/2$ soit -1. La proposition 19.32(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. □

6. Nous attirons votre attention sur le fait que cela n'est en aucun cas une trivialité.

Remarque 19.36.
Vous avez déjà sans doute vu la démonstration de $\cos(30^\circ) = 1/2$ à partir de la figure 19.4. 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 19.151, laquelle se base sur le lemme 19.135 qui, elle-même, se base sur la proposition 19.58.

Bref, à notre niveau, nous sommes encore loin de pouvoir faire des raisonnements trigonométriques sur base de géométrie dans les triangles.

Exemple 19.37
Développer la fonction cos autour de $x = \frac{\pi}{3}$. Utiliser la valeur remarquable du lemme 19.35. Nous développons autour de $h = 0$ la fonction $\cos(\frac{\pi}{3} + h)$:

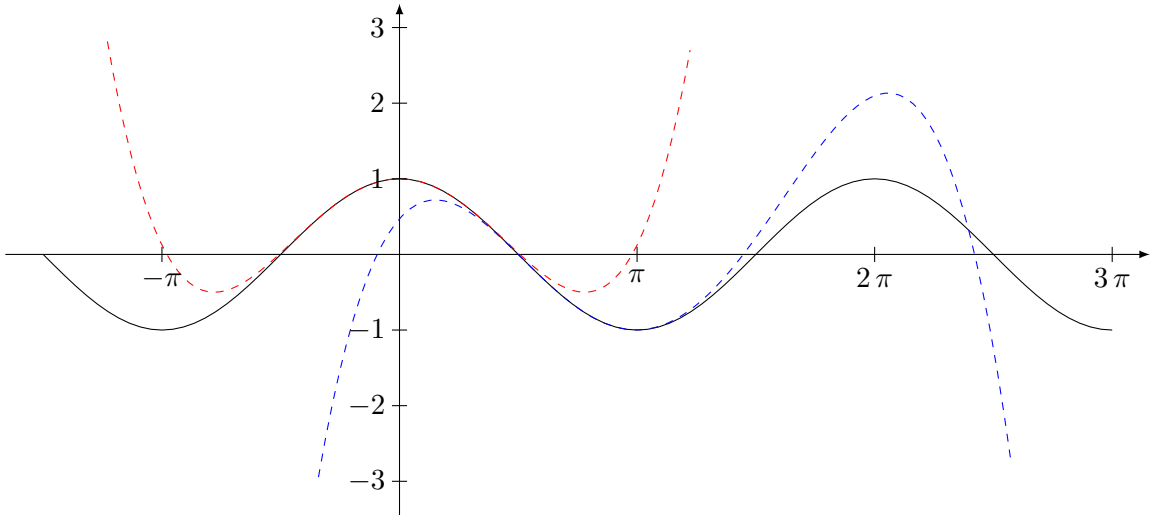
$$\cos\left(\frac{\pi}{3} + h\right) \sim \cos\left(\frac{\pi}{3}\right) + h\cos'(\frac{\pi}{3}) + \frac{h^2}{2}\cos''\left(\frac{\pi}{3}\right) = \frac{1}{2} - \frac{\sqrt{3}}{2}h - \frac{1}{4}h^2. \tag{19.106}$$

Il est aussi possible d'écrire cela en notant $x = x_0 + h$, c'est-à-dire en remplaçant h par $x - \frac{\pi}{3}$:

$$\cos(x) \sim \frac{1}{2} - \frac{\sqrt{3}}{2}(x - \frac{\pi}{3}) - \frac{1}{4}(x - \frac{\pi}{3})^2. \tag{19.107}$$

△

Pour donner une idée nous avons dessiné sur le graphe suivant la fonction sinus et ses développements d'ordre 4 autour de zéro et autour de $3\pi/4$.



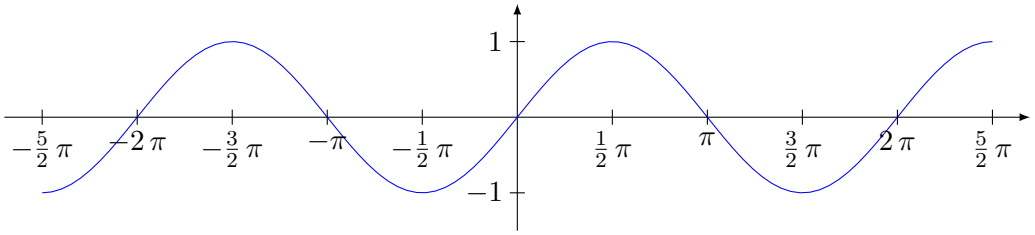
Voici un tableau qui rappelle les valeurs à retenir pour les fonctions sinus, cosinus et tangente.

x	$\sin(x)$	$\cos(x)$	$\tan(x)$
0	0	1	0
$\pi/6$	$1/2$	$\sqrt{3}/2$	$\sqrt{3}/3$
$\pi/6$	$1/2$	$\sqrt{3}/2$	$\sqrt{3}/3$
$\pi/4$	$\sqrt{2}/2$	$\sqrt{2}/2$	1
$\pi/3$	$\sqrt{3}/2$	$1/2$	$\sqrt{3}$
$\pi/2$	1	0	N.D.

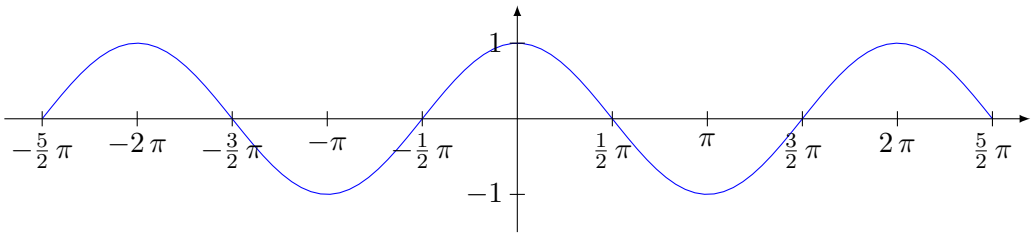
(19.108)

où « N.D. » signifie « non défini ».

Rappelons le graphe de la fonction sinus :



celui de la fonction cosinus :



Lemme 19.38.
Pour toute valeur de $x \in \mathbb{R}$ on a $|\sin(x)| \leqslant |x|$.

Démonstration. Nous séparons des cas en fonction des valeurs.

— Si $0 \leqslant x \leqslant \pi/2$ alors le sinus de x est la longueur du cathète verticale du triangle rectangle de sommets $O = (0,0)$, $A = (\cos(x), \sin(x))$ et $B = (\cos(x), 0)$. Le triangle de sommets A, B et $C = (1,0)$ est aussi rectangle et nous savons que chacun des cathètes ne peut pas être plus long que l'hypoténuse. Donc $\sin(x)$ est inférieur à la longueur du segment AC. À son tour le segment AC ne peut pas être plus long que l'arc de cercle $\widehat{A0C}$, car le chemin le plus court entre deux points du plan est toujours donné par un morceau de droite. La longueur de l'arc de cercle $\frown AC$ est *par définition* la mesure en radiants de l'angle $\widehat{AOC}$, qui est x et on a l'inégalité $\sin(x) \leqslant x$.

— Si $-\pi/2 \leqslant x \leqslant 0$ le même raisonnement que au point précédent permet de conclure que $\sin(x) \leqslant |x|$.

— Nous savons par ailleurs que la fonction sinus prend ses valeurs dans l'intervalle $[-1,1]$ et donc pour tout x tel que $|x| \geqslant \pi/2 \equiv 1,57\ldots$ on a forcement $|\sin(x)| \leqslant |x|$.

□

$$\begin{cases} x = r\cos\theta & (19.109a) \\ y = r\sin\theta & (19.109b) \end{cases}$$

avec $r \in]0,\infty[$ et $\theta \in [0, 2\pi[$. Le jacobien vaut r.

19.4 Très modeste approximation de π

Nous sommes en droit de vouloir une valeur approchée de π.

Lemme 19.39.
Nous avons l'approximation numérique

$$2\sqrt{2} < \pi < 4. \tag{19.110}$$

Démonstration. Grace au lemme 19.34 nous savons que la fonction sin passe de 0 à $\sqrt{2}/2$ sur un intervalle de taille $\pi/4$ avec une dérivé majorée par 1. Par conséquent

$$\frac{\pi}{4} > \frac{\sqrt{2}}{2} \tag{19.111}$$

et donc [7]

$$\pi > 2\sqrt{2} \simeq 2.82 \tag{19.112}$$

De plus la fonction sin passe de 0 à $\sqrt{2}/2$ sur un intervalle de taille $\pi/4$ avec une dérivée majorée par $\sqrt{2}/2$, donc

$$\frac{\pi}{4} < \frac{\sqrt{2}/2}{\sqrt{2}/2}, \tag{19.113}$$

ce qui donne

$$\pi < 4. \tag{19.114}$$

□

Pour avoir une meilleur approximation de π, nous pouvons remarquer que $\pi \in]2.82, 4[$, et que cet intervalle est suffisamment petit pour ne pas recouvrir l'intervalle correspondant pour 2π. L'équation $\cos(x) = -1$ possède donc une unique solution dans cet intervalle (et cette solution est π). Nous pouvons donc faire une dichotomie pour trouver la valeur de π, pourvu que nous ayons une façon d'évaluer des valeurs de $\cos(x)$ de façon pas trop ridicule.

19.5 Cercle trigonométriques

Proposition 19.40 ([317]).
Soient des fonctions $f, g\colon I \to \mathbb{R}$ de classe C^1 sur l'ouvert I de $\mathbb{R}$ telles que $f^2 + g^2 = 1$. Soit $t_0 \in I$ et θ_0 tel que $f(t_0) = \cos(\theta_0)$ et $g(t_0) = \sin(\theta_0)$.

Alors il existe une unique fonction continue $\theta\colon I \to \mathbb{R}$ telle que

$$\begin{cases} \theta(t_0) = \theta_0 & \text{(19.115a)} \\ f = \cos\circ\theta & \text{(19.115b)} \\ g = \sin\circ\theta. & \text{(19.115c)} \end{cases}$$

Démonstration. Nous commençons par l'existence, en passant par les nombres complexes. Soit $h\colon I \to \mathbb{C}$ définie par $h = f + ig$. Nous avons $h\bar{h} = 1$ et nous définissons

$$\theta(t) = \theta_0 - i\int_{t_0}^{t} h'(s)\overline{h(s)}ds. \tag{19.116}$$

Cette intégrale existe pour tout t parce que les fonctions f et g étant de classe C^∞, elles sont bornées sur le compact $[t_0, t]$. De plus θ est une fonction continue parce que c'est une primitive (proposition 15.232) [8].

La dérivée de θ est la fonction $s \mapsto -ih'(s)\overline{h(s)}$.

Utilisant la formule du lemme 19.25 sur la forme trigonométrique des nombres complexes, nous calculons :

$$\frac{d}{dt}\Big[he^{-i\theta}\Big]_{t=0} = e^{-i\theta}(h' - h\theta') = e^{-i\theta}(h' - ih(-i)h'\bar{h}) = 0. \tag{19.117}$$

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

$$h(t_0)e^{-i\theta(t_0)} = c \tag{19.118}$$

donne immédiatement $c = 1$, ou encore $e^{i\theta(t)} = h(t)$, c'est-à-dire que

$$f + ig = \cos\circ\theta + i\sin\circ\theta, \tag{19.119}$$

ce qu'il fallait pour l'existence.

7. Sérieusement, êtes vous capables de trouver une approximation de $\sqrt{2}$ en ne vous basant que sur des choses vues jusqu'ici ?

8. En réalité nous appliquons la proposition 11.11 à chacune des parties réelles et imaginaires de la fonction $s \mapsto h'(s)\overline{h(s)}$.

Pour l'unicité nous supposons avoir une autre fonction, α qui satisfait aux exigences. Pour tout $t \in I$ nous avons

$$e^{i\theta(t)} = e^{i\alpha(t)}. \tag{19.120}$$

Il existe donc une fonction $n\colon I \to \mathbb{N}$ telle que $\theta(t) = \alpha(t) + 2n(t)\pi$. Par continuité de θ et α, la fonction n doit être constante, mais vu que $\theta(t_0) = \alpha(t_0)$ nous avons $n = 1$. □

19.5.1 Les fonctions tangente et arc tangente

Définition 19.41.
La fonction ***tangente*** *est :*

$$\tan(x) = \frac{\sin(x)}{\cos(x)} \tag{19.121}$$

où sin *et* cos *sont de la définition 19.16.*

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 19.1.

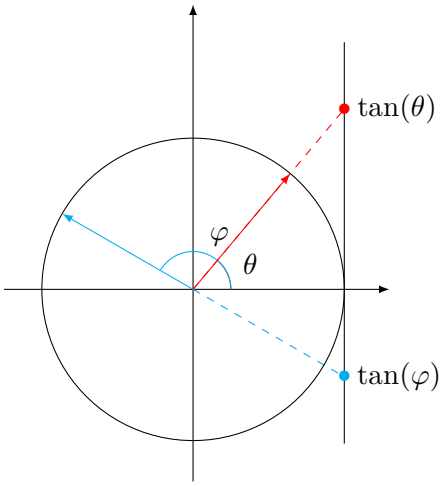


FIGURE 19.1 – Interprétation géométrique de la fonction tangente. La tangente de l'angle θ est positive (et un peu plus grande que 1) tandis que celle de la tangente de l'angle φ est négative.

Proposition 19.42.
La fonction

$$\begin{aligned} \tan\colon]-\frac{\pi}{2}, \frac{\pi}{2}[&\to \mathbb{R} \\ x &\mapsto \tan(x) \end{aligned} \tag{19.122}$$

est une bijection.

Démonstration. Le cosinus ne s'annulant pas sur l'intervalle donné, la fonction est bien définie. Nous avons

$$\lim_{x \to \pi/2^-} \tan(x) = +\infty \tag{19.123}$$

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 [9] dit que la fonction tangente est alors surjective sur $\mathbb{R}$.

Par ailleurs en utilisant les règles de calcul comme la dérivation du quotient 13.114(5) nous trouvons

$$\tan'(x) = \tan^2(x) + 1, \tag{19.124}$$

9. Théorème 13.50.

ce qui nous donne une dérivée partout strictement positive, et donc une fonction strictement croissante et donc injective. □

Le graphe de la fonction tangente est sur la figure 19.2.

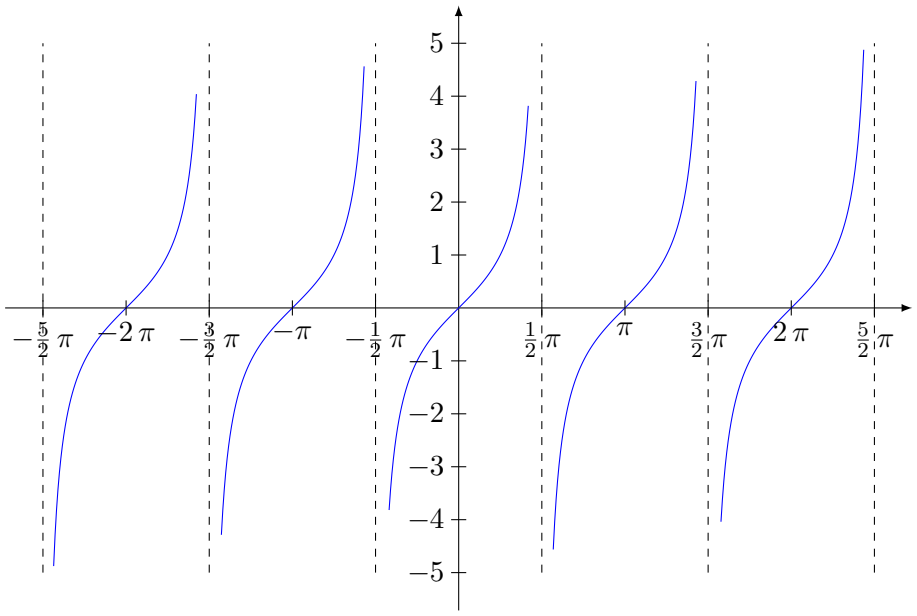


FIGURE 19.2 – Le graphe de la fonction tangente.

En ce qui concerne la bijection réciproque nous avons le théorème suivant.

Théorème 19.43.
La fonction inverse de la tangente,

$$\begin{aligned}\arctan\colon \mathbb{R} &\to \left]-\frac{\pi}{2},\frac{\pi}{2}\right[\\ x &\mapsto \arctan(x)\end{aligned} \tag{19.125}$$

nommée ***arc tangente*** *est*

(1) impaire et strictement croissante sur $\mathbb{R}$.

(2) dérivable sur $\mathbb{R}$ *de dérivée*

$$\arctan'(x) = \frac{1}{1+x^2}. \tag{19.126}$$

Démonstration. Il est immédiatement visible sur son développement de définition (19.52) 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 13.119 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 13.119 et de la dérivée de la tangente : □

Lemme 19.44.
Nous avons les limites

(1) $\lim_{x\to\infty}\arctan(x) = \frac{\pi}{2}$,

(2) $\lim_{x\to-\infty}\arctan(x) = -\frac{\pi}{2}$.

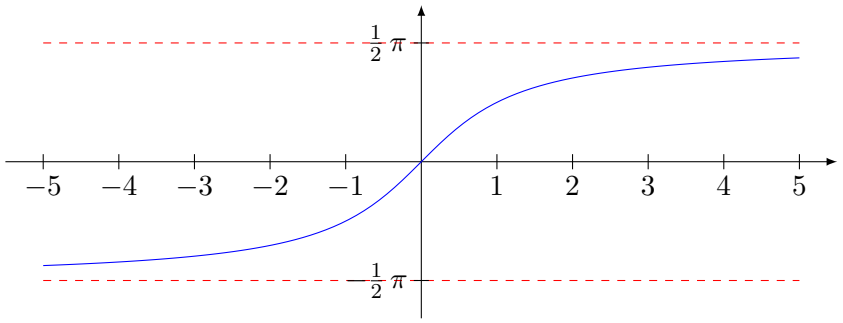
Lemme 19.45.
Nous avons la valeur remarquable

$$\arctan(1/\sqrt{3}) = \frac{\pi}{6}. \tag{19.127}$$

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 :

x	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$
$\arctan(x)$	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$

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 :



19.5.2 La fonction arc sinus

Nous voulons étudier la fonction

$$\begin{aligned} \sin\colon \mathbb{R} &\to [-1,1] \\ x &\mapsto \sin(x) \end{aligned} \tag{19.128}$$

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 13.61, nous devons donc choisir un intervalle à partir duquel la fonction sinus est monotone. Nous choisissons l'intervalle

$$I = [-\frac{\pi}{2}, \frac{\pi}{2}]. \tag{19.129}$$

La fonction

$$\begin{aligned} \sin\colon [-\frac{\pi}{2}, \frac{\pi}{2}] &\to [-1,1] \\ x &\mapsto \sin(x) \end{aligned} \tag{19.130}$$

est une bijection croissante et continue. Nous avons donc le résultat suivant.

Théorème 19.46 (Définition et propriétés de arc sinus)**.**
Nous nommons ***arc sinus*** *la bijection inverse de la fonction* $\sin\colon I \to J$*. La fonction*

$$\begin{aligned} \arcsin\colon [-1,1] &\to [-\frac{\pi}{2}, \frac{\pi}{2}] \\ x &\mapsto \arcsin(x) \end{aligned} \tag{19.131}$$

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 (13.81) s'écrit dans notre cas

$$x = \arcsin\big(\sin(x)\big). \tag{19.132}$$

Nous écrivons d'une part cette équation avec $-x$ au lieu de x :

$$-x = \arcsin\big(\sin(-x)\big) = \arcsin\big(-\sin(x)\big) = \arcsin(-y); \tag{19.133}$$

et d'autre part nous multiplions (19.132) par -1 :

$$-x = -\arcsin\big(\sin(x)\big) = -\arcsin(y). \tag{19.134}$$

En égalisant les valeurs (19.133) et (19.134) nous trouvons

$$\arcsin(-y) = -\arcsin(y), \tag{19.135}$$

ce qui signifie que arcsin est une fonction impaire. □

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 θ_0 dans $[-\frac{\pi}{2},\frac{\pi}{2}]$ pour lequel $\sin(\theta_0) = x_0$. Un tel angle sera forcément unique.

Remarque 19.47.
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.

Exemple 19.48
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 $\arcsin(1) = \frac{\pi}{2}$. △

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.

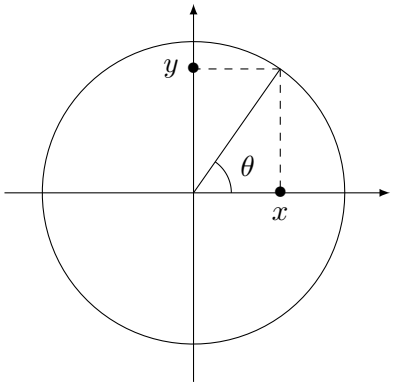
En ce qui concerne la dérivabilité de la fonction arc sinus, en application de la proposition 13.119 elle est dérivable en tout $y = \sin(x)$ tel que $\sin'(x) \neq 0$, c'est-à-dire tel que $\cos(x) \neq 0$. Or $\cos(x) = 0$ pour $x = \pm\frac{\pi}{2}$, ce qui correspond à $y = \sin(\pm\frac{\pi}{2}) = \pm 1$. La fonction arc sinus est donc dérivable sur $]-1,1[$. Nous avons donc la propriété suivante pour la dérivabilité.

Proposition 19.49.
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* (13.275), *qui dans ce cas s'écrit*

$$\arcsin'(y) = \frac{1}{\cos\big(\arcsin(y)\big)} = \frac{1}{\sqrt{1-y^2}}. \tag{19.136}$$

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 19.3.

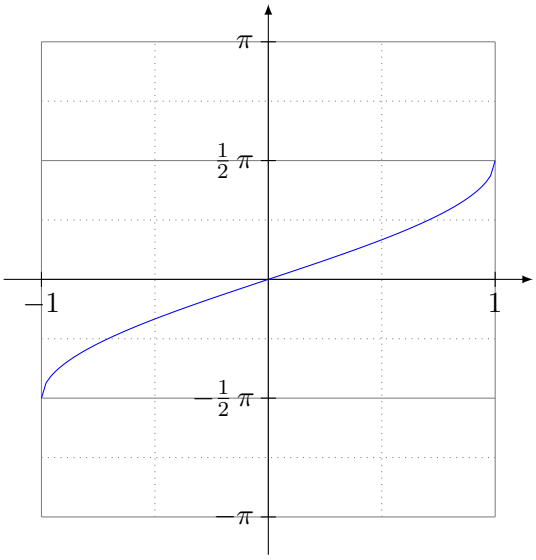


FIGURE 19.3 – Le graphe de la fonction $x \mapsto \arcsin(x)$

19.5.3 La fonction arc cosinus

Nous voulons étudier la fonction

$$\cos\colon \mathbb{R} \to [-1,1] \tag{19.137}$$

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 19.29). 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 :

Proposition-définition 19.50.
Pour définir la fonction arcsinus.

(1) La fonction

$$\cos\colon [0,\pi] \to [-1,1] \tag{19.138}$$

est une bijection continue strictement décroissante.

(2) Sa bijection réciproque est la fonction

$$\arccos\colon [-1,1] \to [0,\pi] \tag{19.139}$$

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

$$\arccos'(y) = \frac{1}{-\sin\big(\arccos(y)\big)} = \frac{-1}{\sqrt{1-y^2}}. \tag{19.140}$$

Démonstration. La fonction cosinus est continue et même de classe C^∞ par la proposition 19.16. Elle est strictement décroissant parce que sa dérivée $(-\sin)$ y est strictement positive (strictement à dans l'intérieur du domaine).

Le fait que arc cosinus soit une bijection continue strictement monotone est dans le théorème de la bijection 13.61. La dérivabilité et la formule sont de la proposition 13.119. □

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 :

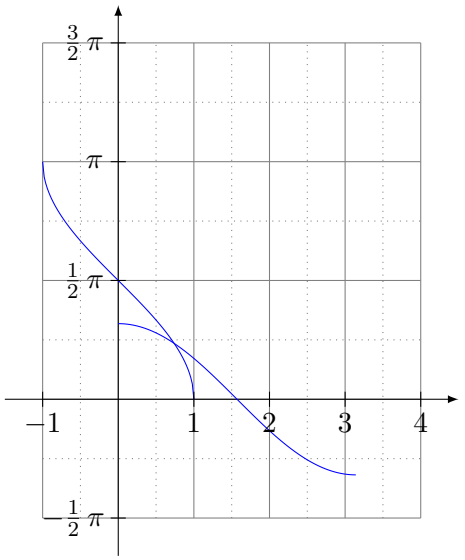
x	-1	$-\frac{\sqrt{3}}{2}$	$-\frac{\sqrt{2}}{2}$	$-\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
$\arccos(x)$	π	$\frac{5\pi}{6}$	$\frac{3}{4}\pi$	$\frac{2}{3}\pi$	$\frac{1}{2}\pi$	$\frac{\pi}{3}$	$\frac{1}{4}\pi$	$\frac{1}{6}\pi$	0

Remarque 19.51.
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 19.52
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 $\theta = \frac{2\pi}{3}$. △

En ce qui concerne la représentation graphique, il suffit de tracer la fonction cosinus entre 0 et π puis de prendre le symétrique par rapport à la droite $y = x$.



19.5.4 Une meilleure approximation de π

Nous avions laissé le nombre π avec l'approximation assez minable de $2\sqrt{2} < \pi < 4$ en le lemme 19.39. Nous pouvons maintenant faire nettement mieux.

Le lemme 19.45 donne

$$\arctan(1/\sqrt{3}) = \pi/6 \tag{19.141}$$

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 calcules peut-être fastidieux, mais comme un gars l'a fait dès l'an 1424[318] pour trouver 16 décimales correctes, nous faisons comme si c'était facile.

Pour trouver le développement en série de Taylor (théorème 13.359) de arc tangente autour de $x = 0$, il faut partir de la formule (19.126) et sans doute pas mal calculer et faire une récurrence [10]. Le résultat est :

$$\arctan(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1}, \tag{19.142}$$

valable pour $x \in]-1, 1[$. Avec cela nous avons

$$\arctan(\frac{1}{\sqrt{3}}) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)3^k} \times \frac{1}{\sqrt{3}} = \frac{\pi}{6}, \tag{19.143}$$

et donc

$$\pi = \frac{6}{\sqrt{3}} \sum_{k=0}^{\infty} \frac{(-1)k}{(2k+1)3^k}. \tag{19.144}$$

Pour donner une idée du fait que ça fonctionne pas mal, voici le calcul pour quelques termes :

```
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: 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}.

Problèmes et choses à faire

Pour bien faire, il faudrait étudier le reste et donner un encadrement.

19.5.5 Forme polaire ou trigonométrique des nombres complexes

Un nombre complexe étant représenté par deux nombres, on peut le représenter dans un plan appelé « plan de Gauss ». La plupart des opérations sur les nombres complexes ont leur interprétation géométrique dans ce plan.

Dans le plan de Gauss, le module d'un complexe z représente la distance entre 0 et z. On appelle *argument* de z (noté $\arg z$) l'angle (déterminé à 2π près) entre le demi-axe des réels positifs et la demi-droite qui part de 0 et passe par z. Le module et l'argument d'un complexe permettent de déterminer univoquement ce complexe puisqu'on a la formule

$$z = a + bi = |z| \left(\cos(\arg(z)) + i \sin(\arg(z))\right)$$

L'argument de z se détermine via les formules

$$\frac{a}{|z|} = \cos(\arg(z)) \quad \frac{b}{|z|} = \sin(\arg(z))$$

10. Je n'ai pas fait le calcul, merci de me faire savoir si il y a une astuce.

ou encore par la formule

$$\frac{b}{a} = \tan(\arg(z)) \quad \text{en vérifiant le quadrant.}$$

La vérification du quadrant vient de ce que la tangente ne détermine l'angle qu'à π près.

19.5.6 Angle entre deux vecteurs

Proposition-définition 19.53.
Soient des vecteurs $X, Y \in \mathbb{R}^2$. Il existe un unique $\theta \in [0, \pi]$ tel que

$$\cos(\theta) = \frac{X \cdot Y}{\|X\|\|Y\|}. \tag{19.145}$$

Ce réel est appelé ***angle*** *entre X et Y.*

Démonstration. Si a et b sont des réels, l'inégalité $|a| \leqslant b$ peut se développer en une double inégalité

$$-b \leqslant a \leqslant b. \tag{19.146}$$

L'inégalité de Cauchy-Schwarz (11.9) devient alors

$$-\|X\|\|Y\| \leqslant X \cdot Y \leqslant \|X\|\|Y\|. \tag{19.147}$$

Si $X \neq 0$ et $Y \neq 0$, nous en déduisons

$$-1 \leqslant \frac{X \cdot Y}{\|X\|\|Y\|} \leqslant 1. \tag{19.148}$$

Il existe donc par la proposition 19.50 un angle $\theta \in [0, \pi]$ tel que

$$\cos(\theta) = \frac{X \cdot Y}{\|X\|\|Y\|}. \tag{19.149}$$

□

19.54.
Certains n'hésitent pas à écrire la formule

$$X \cdot Y = \|X\|\|Y\| \cos(\theta). \tag{19.150}$$

comme une définition du produit scalaire. C'est ce qui arrive lorsqu'on défini les fonctions trigonométriques à partir de relations dans les triangles rectangles.

Notez que les angles entre deux vecteurs sont toujours plus petits ou égaux à 180°.

La longueur de la projection du point P sur la droite horizontale va naturellement être égale à $\cos(\theta)$. En effet, si nous notons X un vecteur horizontal de norme 1, cette projection est donné par $P \cdot X$. Mais en reprenant l'équation (19.150), nous voyons que

$$P \cdot X = \|P\|\|X\| \cos(\theta), \tag{19.151}$$

tandis qu'ici nous avons $\|P\| = \|X\| = 1$.

Nous appelons $\sin(\theta)$ la longueur de la projection sur l'axe vertical.

Quelques dessins nous convainquent que

$$\begin{aligned} \sin(\theta + 2\pi) &= \sin(\theta) \quad & \cos(\theta + 2\pi) &= \sin(\theta), \\ \sin(\theta + \frac{\pi}{2}) &= \cos(\theta) \quad & \cos(\theta + \frac{\pi}{2}) &= -\sin(\theta), \\ \sin(\pi - \theta) &= \sin(\theta) \quad & \cos(\pi - \theta) &= -\cos(\theta). \end{aligned} \tag{19.152}$$

Le théorème de Pythagore nous montre aussi l'importante relation

$$\sin^2(\theta) + \cos^2(\theta) = 1. \tag{19.153}$$

Quelques valeurs remarquables pour les sinus et cosinus :

$$\begin{aligned} &\sin 0 = 0, \quad \sin\frac{\pi}{6} = \frac{1}{2}, \quad \sin\frac{\pi}{4} = \frac{\sqrt{2}}{2}, \quad \sin\frac{\pi}{3} = \frac{\sqrt{3}}{2}, \quad \sin\frac{\pi}{2} = 1, \quad \sin\pi = 0 \\ &\cos 0 = 1, \quad \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}, \quad \cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}, \quad \cos\frac{\pi}{3} = \frac{1}{2}, \quad \cos\frac{\pi}{2} = 0, \quad \cos\pi = -1 \end{aligned} \tag{19.154}$$

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 19.4.

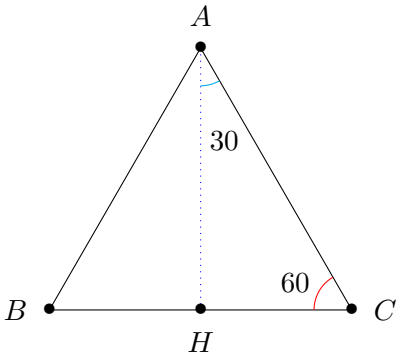


FIGURE 19.4 – Un triangle équilatéral de côté 1.

19.5.7 Aire du parallélogramme

Remarque 19.55.
Le nombre $\|a\|\|b\|\sin(\theta)$ est l'aire du parallélogramme formé par les vecteurs a et b, comme cela se voit sur la figure 19.5.

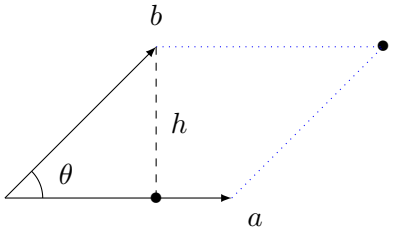


FIGURE 19.5 – Calculer l'aire d'un parallélogramme.

Proposition 19.56.
Nous avons

$$\|a \times b\| = \|a\|\|b\|\sin(\theta) \tag{19.155}$$

où $\theta \in [0.\pi]$ est l'angle formé par a et b.

Démonstration. En utilisant la décomposition du produit vectoriel, nous avons

$$\begin{aligned}\|a\times b\|^2 &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2 \\ &= (a_2b_3-b_2a_3)^2+(a_1b_3-a_3b_1)^2+(a_1b_2-a_2b_1)^2 \\ &= (a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)-(a_1b_1+a_2b_2+a_3b_3)^2 \\ &= \|a\|^2\|b\|^2-(a\cdot b)^2 \\ &= \|a\|^2\|b\|^2-\|a\|^2\|b\|^2\cos^2(\theta) \\ &= \|a\|^2\|b\|^2\big(1-\cos^2(\theta)\big) \\ &= \|a\|^2\|b\|^2\sin^2(\theta).\end{aligned} \tag{19.156}$$

D'où le résultat. Nous avons utilisé la formule de la définition (19.53) donnant l'angle en fonction du produit scalaire. □

19.57.
Si les vecteurs a, b et c ne sont pas coplanaires, alors la valeur absolue du produit mixte (voir équation (11.75)) $a\cdot(b\times c)$ donne le volume du parallélépipède construit sur les vecteurs a, b et c.

En effet si φ 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 [11]

$$V=\|b\times c\|\|a\|\cos(\varphi). \tag{19.157}$$

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 (11.75).

La valeur absolue du déterminant

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \tag{19.158}$$

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 19.55, l'aire de ce parallélogramme est donnée par la norme du produit vectoriel

$$\begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix}\times\begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}=\begin{vmatrix} e_x & e_y & e_z \\ a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \end{vmatrix}=\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}e_z, \tag{19.159}$$

donc la norme $\|a\times b\|$ est bien donnée par la valeur absolue du déterminant (19.158).

19.6 Paramétrisation du cercle

Nous allons parler de paramértisation 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

$$\begin{aligned}\gamma\colon [0,2\pi] &\to \mathbb{R}^2 \\ t &\mapsto \big(\cos(t),\sin(t)\big)\end{aligned} \tag{19.160}$$

11. 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 son développement (définition 19.16).

et

$$\begin{aligned}\varphi\colon [0,2\pi[&\to \mathbb{C}\\ t &\mapsto e^{it}.\end{aligned} \tag{19.161}$$

C'est évidemment la formule $e^{ti} = \cos(t) + i\sin(t)$ (lemme 19.25) qui permet de transformer γ en φ et inversement. De plus $\mathbb{R}^2$ et $\mathbb{C}$ sont isomorphes en tant qu'espaces vectoriels normés (et aussi donc topologiques).

Nous allons voir deux choses à propos de cette application :

— Elle est continue, mais son inverse n'est pas continue. En considérant seulement la restriction $\varphi\colon]0,2\pi[\to S^2\backslash\{(1,0)\}$ nous avons un difféomorphisme, et donc une possibilité de changement de variables dans l'intégrale (théorème 15.252).

Le fait qu'il manque un point est sans importante parce que nous n'allons considérer que la mesure de Lebesgue ou des variations simples autour de la mesure de Lebesgue.

— La fonction $\varphi\colon [0,2\pi[\to S^1$ est une bijection borélienne d'inverse borélien [12]. Donc nous pouvons transposer toute la théorie de la mesure de S^1 à $[0,2\pi[$ sans « triche ».

Tout cela pour dire que nous allons donner un tas de justifications pour écrire des égalités du type

$$\int_{S^1} f = \int_0^{2\pi} f\circ\varphi. \tag{19.162}$$

19.6.1 Bijection continue

Proposition 19.58.
L'application

$$\begin{aligned}\gamma\colon [0,2\pi[&\to S^1\subset\mathbb{R}^2\\ t &\mapsto \big(\cos(t),\sin(t)\big)\end{aligned} \tag{19.163}$$

est une bijection continue.

Démonstration. La continuité découle de la continuité des composantes. Le fait que l'image de γ soit dans S^1 découle immédiatement du fait que $\sin^2+\cos^2=1$.

Pour la bijection, il faut injectif et surjectif.

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 $\sin(x_1)=\sin(x_2)$, il existe par le théorème de Rolle 13.126 un élément $c\in]x_1,x_2[$ tel que $\sin'(c)=0$, c'est-à-dire $\cos(c)=0$. Cela contredirait la proposition 19.32(10) à moins que $x_1=x_2$.

Surjectif Soient x,y tels que $x^2+y^2=1$. Supposons pour varier les plaisirs que $x<0$ et $y>0$. Vu que la fonction cos va de 0 à -1 lorsque x va de $\pi/2$ à π, le théorème des valeurs intermédiaires donne $t\in[\pi/2,\pi]$ tel que $\cos(t)=x$. Pour cette valeur de x nous avons

$$\cos^2(x)+\sin^2(x)=1, \tag{19.164}$$

et donc $\sin^2(x)=y^2$, ce qui donne $\sin(x)=\pm y$. Mais pour $x\in[\pi/2,\pi]$ nous avons $\sin(t)>0$. Par conséquent $\sin(t)=y$.

□

Exemple 19.59
L'application

$$\begin{aligned}\varphi\colon]0,2\pi[&\to S^1\\ x &\mapsto \begin{pmatrix}\cos(x)\\ \sin(x)\end{pmatrix}\end{aligned} \tag{19.165}$$

12. Proposition 19.63.

est un continue par la proposition 19.58. Vu que $]0,2\pi[$ est connexe (proposition 8.34) la proposition 7.84 implique que le cercle privé d'un point est connexe. △

Allez... Dans l'intro nous avions dit que nous n'allions pas faire explicitement les isomorphismes. Faisons-le quand même une fois, mais c'est bien parce que c'est vous hein.

Proposition 19.60.
L'application

$$\begin{aligned} f\colon [0,2\pi[&\to S^1 \\ x &\mapsto e^{ix} \end{aligned} \tag{19.166}$$

est une bijection. Ici, S^1 est l'ensemble des nombres complexes de norme 1.

Démonstration. Nous savons que

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\to \mathbb{C} \\ (x,y) &\mapsto x+iy \end{aligned} \tag{19.167}$$

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

$$\begin{aligned} \gamma\colon [0,2\pi[&\to S^1 \subset \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \end{aligned} \tag{19.168}$$

qui est une bijection continue (c'est la proposition 19.58). Et enfin le lemme 19.25 nous donne $e^{ix} = \cos(x) + i\sin(x)$.

Avec tout ça, l'application $\varphi^{-1} \circ f\colon [0,2\pi[\to S^1$ est une bijection continue. Et comme φ l'est également, f est une bijection continue. □

La proposition suivante donne les coordonnées polaires sur $\mathbb{C}$. La régularité est l'objet du théorème 19.190 (à part le fait que ce dernier parle de $\mathbb{R}^2$ et non de $\mathbb{C}$).

Proposition 19.61.
Pour tout nombre complexe z, il existe un unique $\theta \in [0,2\pi[$ tel que

$$z = |z|e^{i\theta}. \tag{19.169}$$

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 19.60).

Pour ce θ nous avons $z = |z|e^{i\theta}$. □

Bien entendu, le θ est unique dans $[0,2\pi[$, mais il n'est pas du tout unique dans $\mathbb{R}$.

19.6.2 Inverse

Nous pouvons écrire un inverse de la fonction φ grâce à la fonction arc tangente introduite au théorème 19.43. La fonction que nous écrivons à présent est la fonction $\arg_{0-}$ définie par (27.280). Elle n'est pas exactement la fonction argument définie par (27.242).

Nous avons :

$$\varphi^{-1}\colon S^1 \to [0,2\pi[$$

$$x+iy \mapsto \begin{cases} \operatorname{arctg}(y/x) & \text{si } x>0, y\geqslant 0 \\ \frac{\pi}{2} & \text{si } (x,y)=(0,1) \\ \pi - \operatorname{arctg}(-y/x) & \text{si } x<0, y\geqslant 0 \\ \pi + \operatorname{arctg}(y/x) & \text{si } x<0, y<0 \\ 2\pi - \operatorname{arctg}(-y/x) & \text{si } x>0, y<0 \end{cases} \tag{19.170}$$

Chacune des branche est continue parce que la fonction arc tangente l'est. Trois des raccords sont également continus grâce aux limites du lemme 19.44.

L'application φ^{-1} n'est cependant pas continue au point $(1,0)$ [13]. C'est l'objet du lemme suivant.

Lemme 19.62.
L'application $\varphi^{-1}\colon S^1 \to [0,2\pi[$ *n'est pas continue en* $(1,0)$. *Mais elle est continue ailleurs. Autrement dit,*

$$\varphi^{-1}\colon S^1\backslash\{(1,0)\} \to]0,2\pi[\tag{19.171}$$

est continue.

Démonstration. En effet, φ^{-1} serait continue si l'image de tout ouvert de $[0,2\pi[$ par φ serait 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 induites 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 :

$$|\cos(t)+i\sin(t)-1|^2 = \big(\cos(t)-1\big)^2+\sin^2(t) = 2\big(1-\cos(t)\big). \tag{19.172}$$

Soit $\delta>0$, et montrons que $B\big((1,0),\delta\big)\cap S^1$ contient des points d'ordonnées négatives. D'abord il existe $\epsilon>0$ tel que pour $t=2\pi-\epsilon$,

$$2\big(1-\cos(t)\big)<\delta. \tag{19.173}$$

Ensuite pour de tels t, nous avons $\sin(t)<0$. Donc les points de S^1 correspondant à $2\pi-\epsilon$ sont dans $S^1\cap B\big((1,0),\delta\big)$.

Bref, l'image de $[0,\epsilon[$ n'est pas un ouvert de S^1. □

19.6.3 Cercle trigonométrique

Le cercle trigonométrique est le cercle de rayon 1 représenté à la figure 19.6. Sa longueur est 2π.

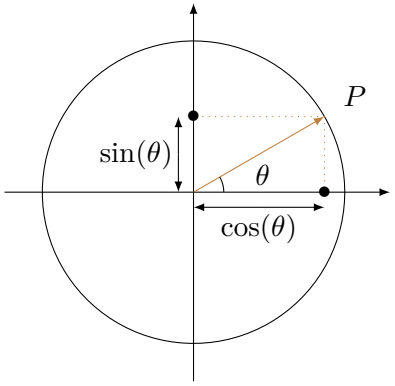


FIGURE 19.6 – Le cercle trigonométrique.

Nous verrons plus tard que la longueur de l'arc de cercle intercepté par un angle θ est égal à θ. Les radians sont donc l'unité d'angle les plus adaptés au calcul de longueurs sur le cercle.

13. Vu que 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 $1\in[0,2\pi[\subset\mathbb{R}$?

19.6.4 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.

Proposition 19.63 ([1]).
L'application φ est borélienne d'inverse borélien, c'est-à-dire

$$\mathcal{B}or(S^1) = \varphi\big(\mathcal{B}or([0,2\pi[)\big). \tag{19.174}$$

Démonstration. L'inclusion $\mathcal{B}\text{or}(S^1) \subset \varphi\big(\mathcal{B}\text{or}([0,2\pi[)\big)$ est la plus simple : si $A \in \mathcal{B}\text{or}(S^1)$, alors $\varphi^{-1}(A) \in \mathcal{B}\text{or}\big([0,2\pi[\big)$ parce que $\varphi\colon [0,2\pi[\to S^1$ est continue et donc borélienne (théorème 15.49).

Pour l'autre inclusion, il faudra faire par étapes.

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 19.62 nous indique que son image par φ est ouverte dans S^1. En particulier, $\varphi(A) \in \mathcal{B}\text{or}(S^1)$.

Ouvert de la forme $[0,\epsilon[$ Nous supposons que ϵ est petit. Disons pour fixer les idées, plus petit que $\pi/2$. Nous avons :

$$\varphi\big([0,\epsilon[\big) = \varphi\big(]0,\epsilon[\big) \cup \varphi\big(\{0\}\big). \tag{19.175}$$

Le premier élément de l'union est un ouvert, et le second un unique point. L'union est un borélien.

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 7.4 couplé au fait que nous sommes dans la topologie induite de $\mathbb{R}$).

Si A est un ouvert contenant zéro, alors

$$A = [0,\epsilon[\cup \big(A \backslash [0,\frac{\epsilon}{2}]\big). \tag{19.176}$$

Nous avons déjà vu que l'image du premier élément de l'union est un borélien. Étant donné que $A\backslash[0,\frac{\epsilon}{2}]$ est un ouvert ne contenant pas zéro, son image est un ouvert. Donc le l'image de A est un borélien.

Pause Nous avons déjà vu que l'image par φ 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

Boréliens Nous utilisons le lemme de transport 15.47 avec l'application φ^{-1} et l'ensemble des ouverts :

$$\varphi\big(\sigma(\mathcal{C})\big) = \sigma\big(\varphi(\mathcal{C})\big) \tag{19.177}$$

où $\mathcal{C}$ est la tribu des ouverts dans $[0,2\pi[$. L'ensemble $\sigma(\mathcal{C})$ est par définition l'ensemble $\mathcal{B}\text{or}\big([0,2\pi[\big)$. D'autre part nous avons vu que l'image d'un ouvert est un borélien : $\varphi(\mathcal{C}) \subset \mathcal{B}\text{or}(S^1)$. Nous avons donc

$$\varphi\big(\mathcal{B}\text{or}([0,2\pi[)\big) = \sigma\big(\varphi(\mathcal{C})\big) \subset \sigma\big(\mathcal{B}\text{or}(S^1)\big) \subset \mathcal{B}\text{or}(S^1). \tag{19.178}$$

La preuve est terminée. □

Proposition 19.64 (Boréliens sur S^1[1]).
Soit la structure usuelle d'espace mesurable $(\mathbb{C}, \mathcal{B}or(\mathbb{C}))$. Nous considérons

- *la tribu $\mathcal{B}or(\mathbb{C})_{S^1}$ induite de la tribu des boréliens de $\mathbb{C}$ vers S^1,*
- *la tribu $\mathcal{B}or(S^1)$ des boréliens de S^1 construite à partir de la topologie induite de $\mathbb{C}$ vers S^1.*

— *la bijection* $\varphi\colon [0,2\pi[\to S^1$,

— *la mesure de Lebesgue sur* $[0,2\pi[$ *(induite de celle sur* $\mathbb{R}$*) et sur* $\mathbb{C}$*, que nous noterons toutes deux* λ.

Alors

(1) Nous avons les expressions

$$\mathcal{B}or(\mathbb{C})_{S^1} = \{A \in \mathcal{B}or(\mathbb{C}) \text{ tel que } A \subset S^1\} \tag{19.179a}$$
$$= \{A \cap S^1 \text{ tel que } A \in \mathcal{B}or(\mathbb{C})\} \tag{19.179b}$$

(2) Nous avons

$$\mathcal{B}or(S^1) = \mathcal{B}or(\mathbb{C})_{S^1} = \varphi\Big(\mathcal{B}or\big([0,2\pi[\big)\Big). \tag{19.180}$$

(3) En définissant $\mu\colon \mathcal{B}or(S^1) \to \mathbb{R}$ *par*

$$\mu(A) = \frac{\lambda\big(\varphi^{-1}(A)\big)}{2\pi}, \tag{19.181}$$

le triple $\big(S^1, \mathcal{B}or(S^1), \mu\big)$ *est un espace mesuré.*

(4) L'espace mesuré $\big(S^1, \mathcal{B}or(S^1), \mu\big)$ *est fini et*

$$\mu(S^1) = 1. \tag{19.182}$$

Démonstration. Point par point.

Pour (1) C'est la proposition 15.7.

Pour (2) La première égalité est le lemme 15.76. Le fait que $\mathcal{B}\text{or}(S^1) = \varphi\Big(\mathcal{B}\text{or}\big([0,2\pi[\big)\Big)$ est déjà la proposition 19.63.

Pour (3) Nous devons d'abord nous assurer que la formule ait un sens. Cela est chose aisée ; si $A \in \mathcal{B}\text{or}(S^1)$, le point (2) nous indique que $\varphi^{-1}(A) \in \mathcal{B}\text{or}\big([0,2\pi[\big)$. Ensuite, nous devons vérifier les deux conditions de la définition 15.20 pour avoir un espace mesuré.

En premier lieu,

$$\mu(\varnothing) = \frac{1}{2\pi}\lambda\big(\varphi^{-1}(\varnothing)\big) = \frac{1}{2\pi}(\varnothing) = 0. \tag{19.183}$$

En en second lieu, si les $A_i \in \mathcal{B}\text{or}(S^1)$ sont disjoints, les $\varphi^{-1}(A_i)$ sont également disjoints parce que φ^{-1} est une bijection. Donc

$$\mu\Big(\bigcup_i A_i\Big) = \frac{1}{2\pi}\lambda\Big(\bigcup_i \varphi^{-1}(A_i)\Big) \tag{19.184a}$$
$$= \frac{1}{2\pi}\sum_i \lambda\big(\varphi^{-1}(A_i)\big) \tag{19.184b}$$
$$= \sum_i \frac{\lambda\big(\varphi^{-1}(A_i)\big)}{2\pi} \tag{19.184c}$$
$$= \sum_i \mu(A_i). \tag{19.184d}$$

D'accord.

Pour (2) En ce qui concerne la mesure de S^1 pour μ nous avons simplement

$$\mu(S^1) = \frac{\lambda\big([0,2\pi[\big)}{2\pi} = 1. \tag{19.185}$$

□

Maintenant que $(S^1, \mathcal{B}or(S^1), \mu)$ est un espace mesuré, nous pouvons compléter la tribu $\mathcal{B}or(S^1)$ pour la mesure μ.

Définition 19.65.
*La **tribu de Lebesgue** sur S^1 est la mesure complétée pour*

$$\big(S^1, \mathcal{B}or(S^1), \mu\big) \tag{19.186}$$

où μ est la mesure définie par la proposition 19.64. Nous notons $\mathcal{L}eb(S^1)$ la tribu et encore μ la mesure.

Proposition 19.66 (Lebesgue sur S^1[1])**.**
Soit la structure d'espace mesuré complet $\big(S^1, \mathcal{L}eb(S^1), \mu\big)$. Nous considérons

— *la tribu $\mathcal{L}eb(\mathbb{C})_{S^1}$ induite de la tribu des boréliens de $\mathbb{C}$ vers S^1,*

— *la bijection $\varphi\colon [0, 2\pi[\to S^1$,*

Alors

(1) La tribu $\mathcal{L}eb(\mathbb{C})_{S^1}$ est la tribu de toutes les parties de S^1.

(2) La tribu $\mathcal{L}eb(S^1)$ est donnée par

$$\mathcal{L}eb(S^1) = \varphi\big(\mathcal{L}eb(\mathbb{R})_{[0,2\pi[}\big) = \varphi\big(\mathcal{L}eb([0, 2\pi[)\big) \tag{19.187}$$

où $\mathcal{L}eb([0, 2\pi[)$ est la tribu sur $[0, 2\pi[$ obtenue par completion de la tribu des boréliens de la topologie induite.

(3) Nous avons l'inclusion stricte

$$\mathcal{L}eb(S^1) \subsetneq \mathcal{L}eb(\mathbb{C})_{S^1}. \tag{19.188}$$

Démonstration. Point par point.

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 λ-négligeable et par conséquent mesurable.

Pour (2) Il s'agit de prouver que

$$\widehat{\mathcal{B}\text{or}(S^1)} = \varphi\big(\widehat{\mathcal{B}\text{or}\big([0, 2\pi[\big)}\big). \tag{19.189}$$

Ce n'est rien d'autre que la proposition 15.71. La seconde partie de l'égalité est la proposition 15.66

Pour (3) Comme indiqué au point (1), la tribu $\mathcal{L}\text{eb}(\mathbb{C})_{S^1}$ est la tribu de toutes les parties de S^1 ; l'incusion 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 $\mathcal{L}\text{eb}(S^1)$.

Soit V non mesurable dans $[0, 2\pi[$ (prenez quelque chose comme l'ensemble de Vitali de l'exemple 15.141). Vu que, par le point (2),

$$\mathcal{L}\text{eb}(S^1) = \varphi\big(\mathcal{L}\text{eb}(\mathbb{R})_{[0,2\pi[}\big), \tag{19.190}$$

la partie $\varphi^{-1}(V)$ ne peut pas être dans $\mathcal{L}\text{eb}(S^1)$.

□

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 28.75 qui donne la structure de

$$L^2\big(S^1, \mathcal{L}\text{eb}(S^1), \mu\big) \tag{19.191}$$

qui sera, sans surprises la même que celle de

$$L^2\big([0, 2\pi[, \mathcal{L}\text{eb}\big([0, 2\pi[\big), \lambda\big). \tag{19.192}$$

19.7 Exemples trigonométriques

Nous mettons ici quelques exemples concernant les fonctions trigonométriques, qui n'ont pas pu être mis dans les chapitres le plus adapté, parce que ces derniers sont plus haut dans la table des matière.

Exemple 19.67
Prouvons que la fonction[14] $f(x) = x\sin(x)$ tend vers zéro lorsque x tend vers 0. D'abord, nous coinçons la fonction entre deux fonctions connues :

$$0 \leqslant |x\sin(x)| = |x||\sin(x)| \leqslant |x|. \tag{19.193}$$

Donc $|x\sin(x)|$ est coincé entre $g(x) = 0$ et $h(x) = |x|$. Ces deux fonctions tendent vers 0 lorsque $x \to 0$, et donc $f(x)$ tend vers zéro. △

19.7.1 Quelque équations trigonométriques

La proposition suivante se voit très facilement sur le cercle trigonométrique, mais il faut le démontrer.

Proposition 19.68 ([1]).
Si $\theta_0 \in [0, 2\pi[$ vérifie $\cos(\theta_0) = x_0$, alors l'ensemble de solutions de l'équation $\cos(\theta) = x_0$ (d'inconnue θ) est

$$\{\theta_0, 2\pi - \theta_0\}. \tag{19.194}$$

Cet ensemble est un singleton si et seulement si $x_0 = \pm 1$.

Démonstration. Commençons par prouver que θ_0 et $2\pi - \theta_0$ sont des solutions. Le nombre θ_0 est solution par hypothèse. En ce qui concerne $2\pi - \theta_0$, il est possible d'utiliser la formule d'addition d'angle (19.72c) :

$$\cos(2\pi - \theta_0) = \cos(2\pi)\cos(\theta_0) + \sin(2\pi)\sin(\theta_0). \tag{19.195}$$

La proposition 19.32(1) nous indique que $\cos(2\pi) = 1$ et $\sin(2\pi) = 0$. Donc l'égalité (19.195) se réduit à $\cos(2\pi - x_0) = \cos(2\pi)$.

Le lemme 19.18 dit que si $\cos(\theta) = x_0$, alors

$$\sin(\theta) = \pm\sqrt{1 - x_0^2}. \tag{19.196}$$

Nous avons donc soit

$$\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix} = \begin{pmatrix}x_0\\ \sqrt{1 - x_0^2}\end{pmatrix}, \tag{19.197}$$

soit

$$\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix} = \begin{pmatrix}x_0\\ -\sqrt{1 - x_0^2}\end{pmatrix}, \tag{19.198}$$

Vu que $\theta \mapsto \big(\cos(\theta), \sin(\theta)\big)$ est une bijection avec S^1 (proposition 19.58), 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 $\big(x_0, \sqrt{1 - x_0^2}\big)$ et $\big(x_0, -\sqrt{1 - x_0^2}\big)$ sont identiques. Cela est le cas si et seulement si $\sqrt{1 - x_0^2} = 0$, c'est à dire si et seulement si $x_0 = \pm 1$. □

14. La définition de la fonction sinus est 19.16.

19.7.2 Développements en série

Proposition 19.69 (Taylor pour cosinus)**.**
Le développement du cosinus est donné par

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!}\cdots \tag{19.199}$$

C'est-à-dire que pout tout $n \in \mathbb{N}$, il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0} \alpha(t) = 0$ et

$$\cos(x) = \sum_{k=0}^{n} \frac{(-1)^{2k}}{(2k)!}x^{2k} + \alpha(x)x^{2n+1}. \tag{19.200}$$

En ce qui concerne le sinus, pour tout n nous avons une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0} \alpha(t) = 0$ et

$$\sin(x) = \sum_{k=0}^{n} \frac{(-1)^{k}x^{2k+1}}{(2k+1)!} + x^{2n+2}\alpha(x). \tag{19.201}$$

Démonstration. Il s'agit d'utiliser la proposition 16.32, en faisant attention à l'ordre. Le fait est que dans (19.200), 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. □

Remarque 19.70.
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 α : 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 une polynôme de degré $2p$, nous avons le développement d'ordre $2p+1$.

(3) Nous aurions pu utiliser les dérivées données dans la proposition 19.17 et les valeurs spéciales (19.56).

Corollaire 19.71.
Il existe une fonction $\alpha\colon \mathbb{R} \to \mathbb{R}$ telle que $\lim_{t\to 0} \alpha(t)/t = 0$ et

$$\sin(x) = x + \alpha(x). \tag{19.202}$$

Nous avons la limite

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1. \tag{19.203}$$

Démonstration. Il s'agit de prendre la formule (19.201) avec $n = 0$. Cela donne tout de suite (19.202). Pour la limite, on divise par x, ce qui donne (pour tout $x \neq 0$)

$$\frac{\sin(x)}{x} = 1 + \frac{\alpha(x)}{x}. \tag{19.204}$$

Et justement la fonction α la la propriété que $\lim_{x\to 0} \alpha(x)/x = 0$. □

Exemple 19.72
Cherchons le développement limité à l'ordre 5 de $\tan(x) = \frac{\sin(x)}{\cos(x)}$. Nous utilisons les développements de la proposition 19.69 :

$$\sin(x) = x - \frac{x^3}{6} + \frac{x^5}{120} + x^5\alpha_1(x) \tag{19.205a}$$

$$\cos(x) = 1 - \frac{x^2}{2} + \frac{x^4}{24} + x^5\alpha_2(x). \tag{19.205b}$$

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 !) :

$$\begin{array}{rl|l}
x - \frac{1}{6}x^3 + \frac{1}{120}x^5 & & 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 \\
-\big(x - \frac{1}{2}x^3 + \frac{1}{24}x^5 \big) & & \hline \\
\hline
\frac{1}{3}x^3 - \frac{1}{30}x^5 & & \\
-\big(\frac{1}{3}x^3 - \frac{1}{6}x^5 + \frac{1}{72}x^7 \big) & & \\
\hline
\frac{2}{15}x^5 - \frac{1}{72}x^7 & & \\
-\big(\frac{2}{15}x^5 - \frac{1}{15}x^7 + \frac{1}{180}x^9 \big) & & \\
\hline
\frac{29}{360}x^7 - \frac{1}{180}x^9 & &
\end{array}$$

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 13.383)

$$\tan(x) = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + x^5\alpha(x). \tag{19.206}$$

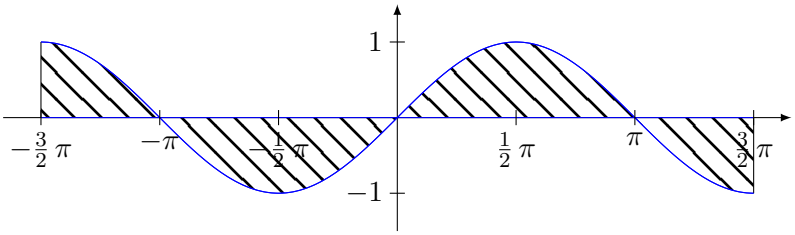
Notons que, vu que le reste ne nous intéresse pas vraiment, nous aurions pu ne pas calculer les coefficients des termes en x^7 et x^8. La dernière soustraction était également inutile. △

19.7.3 Intégration

Exemple 19.73
Comme nous le voyons sur le dessin suivant,

$$\int_{-3\pi/2}^{3\pi/2} \sin(x)\,dx = 0 \tag{19.207}$$

parce que les deux parties bleues s'annulent avec les deux parties rouges (qui sont comptées comme des aires négatives).



△

19.7.4 Changement de variables dans une intégrale

Exemple 19.74
Cet exemple util

Soit V la région trapézoïdale de sommets $(0,-1)$, $(1,0)$, $(2,0)$, $(0,-2)$, comme à la figure 19.7(a). Calculons ensemble l'intégrale double

$$\int_V e^{\frac{x+y}{x-y}}\,dV,$$

avec le changement de variable $\psi(x,y) = (x+y, x-y)$. C'est-à-dire que nous considérons les nouvelles variables

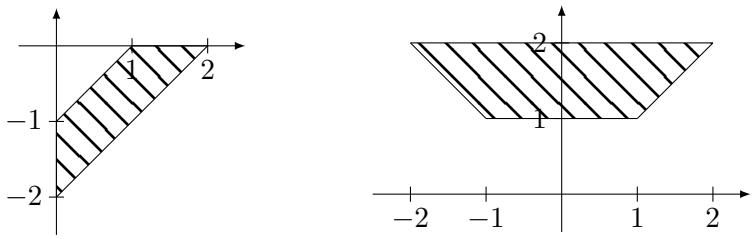
$$\begin{cases} u = x + y & \text{(19.208a)} \\ v = x - y. & \text{(19.208b)} \end{cases}$$

Il faut remarquer d'abord que le changement de variable proposé est dans le mauvais sens. On écrit alors $\phi(u,v) = \psi^{-1}(u,v) = \big((u+v)/2, (u-v)/2\big)$, c'est-à-dire

$$\begin{cases} x = \dfrac{u+v}{2} & (19.209a) \\ y = \dfrac{u-v}{2}. & (19.209b) \end{cases}$$

La région qui correspond à V est U, le trapèze de sommets $(-1,1)$, $(1,1)$, $(2,2)$ et $(-2,2)$, qu'on voit sur la figure 19.7(b) et qu'on décrit par

$$U = \{(u,v) \in \mathbb{R}^2 \,|\, 1 \leqslant v \leqslant 2,\ -v \leqslant u \leqslant v\}.$$



(a) La région V (b) La région $U = \phi^{-1}(V)$

FIGURE 19.7 – Avant et après le changement de variables

On observe que U est une région du premier type tandis que V n'est pas du premier ou du deuxième type. Le déterminant de la matrice jacobienne de ψ^{-1} est $J_{\psi^{-1}}$,

$$J_{\psi^{-1}}(u,v) = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}. \tag{19.210}$$

On a alors, en utilisant le fait que $F(x) = ae^{x/a}$ est une primitive de $f(x) = e^{x/a}$ (proposition 16.54) ainsi que le théorème fondamental de l'analyse (théorème 15.233),

$$\int_V e^{\frac{x+y}{x-y}}\,dV = \int_U e^{\frac{u}{v}}\,\frac{1}{2}\,dV = \int_1^2 \int_{-v}^{v} e^{\frac{u}{v}}\,\frac{1}{2}\,du\,dv = \frac{3}{4}(e - e^{-1}).$$

△

19.8 Isométries dans $\mathbb{R}^n$

Définition 19.75.
*Un **hyperplan** de $\mathbb{R}^n$ est un sous-espace affine de dimension $n-1$.*

Lemme-définition 19.76.
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) $x - y \perp H$,

(2) Le segment $[x,y]$ coupe H en son milieu.

*La **réflexion** σ_H est l'application $\sigma_H \colon \mathbb{R}^n \to \mathbb{R}^n$ 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 cela 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

$$H = H_0 + \lambda e_n \tag{19.211}$$

pour un certain λ.

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 droit à être dans H est le point tel que $x_n e_n + te_n = \lambda e_n$, c'est-à-dire $t = -x_n$. L'unique point y sur cette droite à être tel que $[x, y]$ coupe H en son milieu est celui qui correspond à $t = -2x_n$. □

Notons au passage que cette preuve donne une formule pour σ_H :

$$\sigma_H(x) = \sum_{i=1}^{n-1} x_i e_i - x_n e_n. \tag{19.212}$$

Il s'agit donc de changer le signe de la composante perpendiculaire à H.

Lemme 19.77.
Dans cette même base si H_0 est l'hyperplan parallèle à H et passant par l'origine, nous écrivons $H = H_0 + \lambda e_n$ pour un certain λ. Alors

$$\sigma_H = \sigma_{H_0} + 2\lambda e_n. \tag{19.213}$$

Démonstration. Un élément $x \in \mathbb{R}^n$ peut être décomposé dans la base adéquate en $x = x_H + x_n e_n$. Nous savons de la formule (19.212) que

$$\sigma_H(x) = x_H - x_n e_n. \tag{19.214}$$

Mais vu que $\sigma_{H_0}(x_H) = x_H - 2\lambda e_n$ nous avons

$$\sigma_{H_0}(x) + 2\lambda e_n = \sigma_{H_0}(x_H + x_n e_n) + 2\lambda e_N = x_H - 2\lambda e_n - x_n e_n + 2\lambda e_n = x_H - x_n e_n. \tag{19.215}$$

□

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).

Lemme 19.78 ([313])**.**
Soient deux points distincts $x_0, y_0 \in \mathbb{R}^n$ l'ensemble $H \subset \mathbb{R}^n$ donné par

$$H = \{x \in \mathbb{R}^n \text{ tel que } d(x, x_0) = d(x, y_0)\}. \tag{19.216}$$

Alors H est l'hyperplan orthogonal au vecteur $v = y_0 - x_0$ et H passe par le milieu du segment $[x_0, y_0]$.

Démonstration. Nous savons que

$$d(x, x_0)^2 = \langle x - x_0, x - x_0 \rangle = \|x\|^2 + \|x_0\|^2 - 2\langle x, x_0 \rangle, \tag{19.217}$$

ou encore

$$\|x_0\|^2 - \|y_0\|^2 = 2\langle x, x_0 - y_0 \rangle. \tag{19.218}$$

En posant $v = y_0 - x_0$ et en considérant la forme linéaire

$$\begin{aligned} \beta \colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \langle x, v \rangle, \end{aligned} \tag{19.219}$$

Nous avons $x \in H$ si et seulement si $\beta(x) = \frac{1}{2}\big(\|y_0\|^2 - \|x_0\|^2\big) = \lambda$. En d'autres termes, $H = \beta^{-1}(\lambda)$. Par la proposition 10.23 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 $\dim(H) = \dim\big(\ker(\beta)\big) = n - 1$.

Le fait que H contienne le milieu du segment $[x_0, y_0]$ est par définition. □

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.

Lemme 19.79 ([319]).
Soit un espace euclidien E de dimension n.

(1) Si f est une isométrie de E satisfaisant

$$\dim\big(\mathrm{Fix}(f)\big) = n - k \tag{19.220}$$

alors f peut être écrit comme composition de k réflexions hyperplanes.

(2) Une isométrie de E peut être écrite sous la forme de $\mathrm{rang}(f-\mathrm{Id})$ réflexions, mais pas moins.

(3) Toute isométrie de $\mathbb{R}^n$ 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 sont des reformulations.

Pour (1) Nous faisons une récurrence sur $k \geqslant 0$.

Pour l'initialisation, si $k = 0$ alors $\dim\big(\mathrm{Fix}(f)\big) = n$, 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 $k \geqslant 0$. Soit donc $f \in \mathrm{Isom}(\mathbb{R}^n)$ tel que

$$\dim\big(\mathrm{Fix}(f)\big) = n - (k+1). \tag{19.221}$$

Vu que $k \geqslant 0$, la dimension de $\mathrm{Fix}(f)$ 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

$$H = \{x \in E \text{ tel que } d(x,x_0) = d\big(x, f(x_0)\big)\}. \tag{19.222}$$

Par le lemme 19.78, 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 $g = \sigma_H \circ f$. Vu que $g(x_0) = \sigma_H(f(x_0)) = x_0$, ce x_0 est un point fixe de g. Le fait que $\sigma_H\big(f(x_0)\big) = x_0$ est vraiment la définition de l'hyperplan H.

Nous avons donc

$$x_0 \in \mathrm{Fix}(g) \backslash \mathrm{Fix}(f). \tag{19.223}$$

Mais nous prouvons de plus que $\mathrm{Fix}(f) \subset \mathrm{Fix}(g)$. En effet si $y \in Fix(f)$ alors $y \in H$ parce que

$$d(y,x_0) = d\big(f(y), f(x_0)\big) = d\big(y, f(x_0)\big). \tag{19.224}$$

Vu que $y \in H$ nous avons $y \in \mathrm{Fix}(g)$ parce que

$$g(y) = \sigma_H\big(f(y)\big) = \sigma_H(y) = y. \tag{19.225}$$

Tout cela pour dire que l'ensemble $\mathrm{Fix}(g)$ est *strictement* plus grand que $\mathrm{Fix}(f)$. Et comme ce sont des espaces affines nous pouvons parler de dimension :

$$\dim\big(\mathrm{Fix}(g)\big) > \dim\big(\mathrm{Fix}(f)\big). \tag{19.226}$$

Par hypothèse de récurrence, l'application g peut être écrite comme composition de k réflexions. Donc l'application

$$f = \sigma_H \circ g \tag{19.227}$$

est une composition de $k+1$ réflexions.

Pour (2), existence Ce point est une reformulation du point (1). Le fait est que $\mathrm{Fix}(f) = \ker(f - \mathrm{Id})$ parce que $x \in \mathrm{Fix}(f)$ si et seulement si $f(x) = x$ si et seulement si $(f-\mathrm{Id})x = 0$. Nous utilisons le théorème du rang 4.39 à l'endomorphisme $f - \mathrm{Id}$:

$$\dim\big(\mathrm{Fix}(f)\big) = \dim\big(\ker(f-\mathrm{Id})\big) = \dim(E) - \mathrm{rang}(f-\mathrm{Id}). \tag{19.228}$$

En remplaçant par les valeurs :

$$n-k=n-\operatorname{rang}(f-\mathrm{Id}). \tag{19.229}$$

Or le point (1) donnait f comme composée de $n-k$ réflexions. Donc f est composée de $\operatorname{rang}(f-\mathrm{Id})$ réflexions.

Pour (2), « pas moins » Supposons que $f=\sigma_1\circ\ldots\circ\sigma_r$ où σ_i est la réflexion de l'hyperplan H_i. Nous devons prouver que $r\geqslant\operatorname{rang}(f-\mathrm{Id})$. Nous avons

$$\bigcap_{i=1}^{r}H_i\subset\ker(f-\mathrm{Id}). \tag{19.230}$$

D'autre part, la proposition 11.259 nous donne $\dim\bigcap_i H_i\geqslant n-r$. Donc

$$n-r\leqslant\dim\big(\bigcap_{i=1}^{r}\big)\leqslant\dim\big(\ker(f-\mathrm{Id})\big)=n-\operatorname{rang}(f-\mathrm{Id}). \tag{19.231}$$

Donc $n-r\leqslant n-\operatorname{rang}(f-\mathrm{Id})$ ou encore

$$r\geqslant\operatorname{rang}(f-\mathrm{Id}). \tag{19.232}$$

Pour (3) La première partie de ce théorème n'est rien d'autre que le lemme 19.79 parce que le pire cas est celui où le fixateur de f est réduit à l'ensemble vide, et dans ce cas l'application f est une composition de $n+1$ réflexions.

□

Proposition 19.80.
Un élément de SO(3) *qui fixe deux vecteurs linéairement indépendants est l'identité.*

Démonstration. Soit un élément $A\in\mathrm{SO}(3)$ et deux vecteurs linéairement indépendants $v_1,v_3\in\mathbb{R}^3$ tels que $Av_1=v_1$ et $Av_2=v_2$. Vu que v_1 et v_2 sont linéairement indépendants, le théorème de la base incomplète 4.11 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

$$A=\begin{pmatrix}1&0&a\\0&1&b\\0&0&c\end{pmatrix}. \tag{19.233}$$

Le déterminant de cette matrice est c. Or $\det(A)=1$ parce qu'elle est dans 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 $A=\mathrm{Id}$. □

Corollaire 19.81.
Tout élément de SO(3) *peut être écrit comme composée de deux réflexions.*

Démonstration. Un élément de SO(3) est une isométrie de $\mathbb{R}^3$ parce que si $A\in\mathrm{SO}(3)$ alors [15]

$$\langle Ax,Ay\rangle=\langle A^*Ax,y\rangle=\langle x,y\rangle. \tag{19.234}$$

Donc si le rang de A est k, alors A est la composée de $3-k$ réflexions par le lemme 19.79.

Si $A=\mathrm{Id}$, 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 13.100, 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 $\lambda>0$.

15. Opérateur orthogonal, définition 11.80.

Si v est un vecteur propre correspondant à la valeur propre λ, 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 19.80 nous indique qu'il n'y a pas d'autres vecteurs de $\mathbb{R}^3$ à être fixé par A. Donc $\dim\big(\text{Fix}(A)\big) = 1$ et le lemme 19.79(1) nous s'écrit avec $n = 3$, $k = 2$ et implique que A est la composée de deux réflexions. □

Lemme 19.82.
Soit un hyperplan H et un vecteur v de $\mathbb{R}^n$. Nous avons

$$\tau_v \circ \sigma_H \circ \tau_v^{-1} = \sigma_{\tau_v(H)}. \tag{19.235}$$

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 $\lambda \in \mathbb{R}$,

$$H = H_0 + \lambda e_n \tag{19.236}$$

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

$$\tau_v(H) = H + v = H + v_2 = H_0 + \lambda e_n + v_2. \tag{19.237}$$

Nous pouvons maintenant utiliser le lemme 19.77 pour exprimer la transformation $\sigma_{\tau_v(H)}$:

$$\sigma_{\tau_v(H)}(x) = \sigma_{H_0}(x) + 2\lambda e_n + 2v_2 \tag{19.238}$$

Mais d'autre part,

$$(\tau_v \circ \sigma_H \circ \tau_v^{-1})(x) = v + \sigma_H(x - v) = v + \sigma_{H_0}(x - v) + 2\lambda e_n. \tag{19.239}$$

Vue la décomposition de $v = v_1 + v_2$ nous avons $\sigma_{H_0}(v) = -v_1 + v_2$ et donc

$$(\tau_v \circ \sigma_H \circ \tau_v^{-1})(x) = v + \sigma_{H_0}(x) + v_1 - v_2 + 2\lambda e_n = \sigma_{H_0} + 2v_1 + 2\lambda e_n. \tag{19.240}$$

Les expressions (19.238) et (19.240) coïncident, d'où l'égalité recherchée. □

Théorème 19.83 ([313]).
Une isométrie de $(\mathbb{R}^n, d)$ préserve l'orientation si et seulement si est elle composition d'un nombre pair de réflexions.

Démonstration. Nous définissons

$$\begin{aligned} \epsilon\colon \text{Isom}(\mathbb{R}^n) &\to \{\pm 1\} \\ \tau_v \circ \alpha &\mapsto \det(\alpha) \end{aligned} \tag{19.241}$$

où nous nous référons à la décomposition unique d'un élément de $\text{Isom}(\mathbb{R}^n)$ sous la forme $\tau_v \circ \alpha$ avec $\alpha \in O(n)$ donnée par le théorème 19.1(3).

Le noyau de ϵ est alors la partie

$$\ker(\epsilon) = \mathbb{R}^n \times_{\mathbf{Ad}} \text{SO}(n). \tag{19.242}$$

Une isométrie f préserve l'orientation si et seulement si $\epsilon(f) = 1$. Vu que toutes les isométries sont des compositions de réflexions (première partie), il nous suffit de montrer que $\epsilon(\epsilon_H) = -1$ pour qu'une isométrie préserve l'orientation si et seulement si elle est composition d'un nombre pair de réflexions.

Nous commençons par prouver que pour tout vecteur v, $\epsilon\big(\sigma_H\big) = \epsilon\big(\sigma_{\tau_v(H)}\big)$. Pour cela nous utilisons le lemme 19.82 et le fait que ϵ est un homomorphisme :

$$\epsilon(\sigma_{\tau_v(H)}) = \epsilon(\tau_v)\epsilon(\sigma_H)\epsilon(\tau_v^{-1}) = \epsilon(\sigma_H) \tag{19.243}$$

parce que la partie linéaire d'une translation est l'identité (et donc $\epsilon(\tau_v) = 1$ pour tout v).

Nous avons donc $\epsilon(\sigma_H) = \epsilon(\sigma_{H_0})$. En ce qui concerne σ_{H_0}, dans la base adaptée la matrice est

$$\sigma_{H_0} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & -1 \end{pmatrix}, \tag{19.244}$$

dont le déterminant est -1. □

Pour en savoir plus sur le groupe des isométries, il faut lire le théorème de Cartan-Dieudonné dans [314].

19.9 Groupes finis d'isométries

Définition 19.84.
*Si X est une partie finie de $\mathbb{R}^n$, le **barycentre** de X est le point*

$$B_X = \frac{1}{|X|} \sum_{x \in X} x \tag{19.245}$$

où $|X|$ est le cardinal de X.

Cela est à mettre en relation avec la définition dans le cadre affine 10.27.

Lemme 19.85 ([313])**.**
Les applications affines de $\mathbb{R}^n$ préservent le barycentre[16] *des parties finies.*

Démonstration. Soit une partie finie X de $\mathbb{R}^n$ et une application affine $f \in \mathrm{Aff}(\mathbb{R}^n)$. Nous devons prouver que

$$f(B_X) = B_{f(X)}. \tag{19.246}$$

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 $g \in \mathrm{GL}(n, \mathbb{R})$. Nous vérifions le résultat séparément pour τ_v et pour g.

D'une part,

$$B_{\tau_v(X)} = \frac{1}{|\tau_v(X)|} \sum_{y \in \tau_v(X)} y = \frac{1}{|X|} \sum_{x \in X} (x + v) = B_x + \frac{1}{|X|} \sum_{x \in X} v = B_x + v = \tau_v(B_X). \tag{19.247}$$

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,

$$B_{g(X)} = \frac{1}{|X|} \sum_{x \in X} g(x) = g\Big(\frac{1}{|X|} \sum_{x \in X} x\Big) = g(B_X) \tag{19.248}$$

où nous avons utilisé la linéarité de g dans tous ses retranchements. □

Proposition 19.86.
Points fixes d'un sous-groupe.

(1) Soit H un sous-groupe finie 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 $\mathrm{Isom}(\mathbb{R}^n, d)$ n'acceptant pas de points fixes, alors il est infini.

16. Définition 19.84.

Démonstration. Le groupe H agit sur $\mathbb{R}^n$, et si $x \in \mathbb{R}^n$ nous pouvons considérer son orbite

$$Hx = \{f(x) \text{ tel que } f \in H\}, \tag{19.249}$$

qui est une partie finie de $\mathbb{R}^n$. Considérons son barycentre $v = B_{Hx}$. Soit $f \in H$. Alors

$$\begin{aligned} f(v) &= f(B_{Hx}) && (19.250a)\\ &= B_{f(Hx)} && (19.250b)\\ &= B_{Hx} && (19.250c)\\ &= v, && (19.250d) \end{aligned}$$

Justifications :

— Pour (19.250b), c'est le lemme 19.85.

— Pour (19.250c), c'est le fait que, $f \in H$ étant donné, l'application $g \mapsto fg$ est une bijection de H, donc

$$f(Hx) = \{(fg)(x) \text{ tel que } h \in H\} = \{g(x) \text{ tel que } g \in H\} = Hx. \tag{19.251}$$

Bref, v est fixé par H.

La seconde affirmation n'est rien d'autre que la contraposée de la première. □

Proposition 19.87.
À propos de groupes finis d'isométries.

(1) Tout sous-groupe fini de $\mathrm{Isom}(\mathbb{R}^n)$ *est isomorphe à un sous-groupe fini de* $\mathrm{O}(n)$.

(2) Tout sous-groupe fini de $\mathrm{Isom}^+(\mathbb{R}^n)$ *est isomorphe à un sous-groupe fini de* $\mathrm{SO}(n)$.

Démonstration. Soit H un sous-groupe fini de $\mathrm{Isom}(\mathbb{R}^n)$ et $v \in \mathbb{R}^n$ un élément fixé par H (comme garantit par la proposition 19.86). Nous posons

$$\begin{aligned} \phi\colon H &\to \mathrm{Isom}(\mathbb{R}^n)\\ f &\mapsto \tau_v^{-1} \circ f \circ \tau_v. \end{aligned} \tag{19.252}$$

ϕ est un homomorphisme Les opération du type $\phi = \mathbf{Ad}(\tau_v)$ sont toujours des homomorphismes.

ϕ consiste à extraire la partie linéaire Si $f = \tau_w \circ g$ alors

$$\begin{aligned} \phi(f)(x) &= (\tau_{-v} \circ \tau_w \circ g \circ \tau_v)(x) && (19.253a)\\ &= \tau_{w-v}(g(x) + g(v)) && (19.253b)\\ &= g(x) + g(v) - v + w && (19.253c) \end{aligned}$$

Mais $g(v) + w = f(v)$ et nous savons que $f(v) = v$. Donc il ne reste que $\phi(f)(x) = g(x)$.

ϕ est injective Si $f = \tau_w \circ g$ vérifie $\phi(f) = \mathrm{Id}$, il faut en particulier que $g = \mathrm{Id}$. Mais H est fini et ne peut donc pas contenir de translations non triviales. Donc $w = 0$ et $f = \mathrm{Id}$.

Donc ϕ est une injection à valeur dans les transformations linéaires de $\mathrm{Isom}(\mathbb{R}^n)$. Autrement dit, ϕ est un isomorphisme entre H et son image, laquelle image est dans $\mathrm{O}(n)$.

En ce qui concerne la seconde partie, si $f \in \mathrm{Isom}^+(\mathbb{R}^n)$, alors $\phi(f)$ y est aussi, tout en étant linéaire. Donc $\phi(f) \in \mathrm{SO}(n)$. □

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, ce qui n'est pas possible si H est fini.

19.9.1 Groupe diédral

Proposition 19.88.
Les racines de l'unité dans $\mathbb{C}$, c'est à dire la partie

$$\{e^{2ik\pi/n}, k = 0, \dots, n-1\}, \tag{19.254}$$

forment un polynôme régulier.

19.9.1.1 Définition et générateurs : vue géométrique

Définition 19.89.
*Le **groupe diédral** D_n ($n \geqslant 3$) est le groupe des isométries de $(\mathbb{C}, d)$ laissant invariant l'ensemble*

$$\{e^{2ik\pi/n}, k = 0, \dots, n-1\} \tag{19.255}$$

des racines de l'unité.

19.90.
La proposition 19.88 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 $n \geqslant 3$; et un peu aussi pour une raison technique qui arrivera dans la preuve du lemme 19.91.

Lemme 19.91.
Nous avons

$$D_n \subset O(2, \mathbb{R}). \tag{19.256}$$

Démonstration. Si $f \in D_n$, alors $f(e^{2ik\pi/n})$ doit être l'un des $e^{2ik'\pi/n}$, et vu que f conserve les longueurs dans $\mathbb{C}$, nous devons avoir

$$1 = d(0, e^{2ik\pi/n}) = d\big(f(0), e^{2ik'\pi/n}\big). \tag{19.257}$$

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 $n \geqslant 3$). Par conséquent notre étude du groupe diédral ne doit prendre en compte que les isométries vectorielles de $\mathbb{R}^2$. En d'autres termes

$$D_n \subset O(2, \mathbb{R}). \tag{19.258}$$

□

Proposition 19.92 ([320]).
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[17]. C'est un élément d'ordre 2 qui est dans D_n parce que

$$s\big(e^{2ki\pi/n}\big) = e^{2(n-k)i\pi/n}. \tag{19.259}$$

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 un le groupe d'ordre n des rotations d'angle $2k\pi/n$. □

Proposition 19.93 ([320]).
Si s est la conjugaison complexe et r la rotation d'angle $2\pi/n$, alors $(sr)^2 = \mathrm{Id}$.

Démonstration. Si $z^n = 1$, alors

$$(srsr)z = srse^{2i\pi/n}z = sr\big(e^{-2\pi i/n}\bar{z}\big) = s\bar{z} = z. \tag{19.260}$$

□

17. C'est une réflexion ; la réflexion d'axe $\mathbb{R}$ dans $\mathbb{C}$.

Proposition 19.94 ([320]).
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 $s \circ r^m$.

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

$$r(A_k) = A_{k+1} \tag{19.261a}$$
$$s(A_k) = A_{n-k}. \tag{19.261b}$$

Cette dernière est l'équation (19.259).

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 que un des A_k est fixé par f. Dans ce cas f a deux points fixes : O et A_k et est donc la réflexion d'axe (OA_k). Dans ce cas, nous avons $f = s \circ r^{n-2k}$. En effet

$$s \circ r^{n-2k}(A_k) = s(A_{k+n-2k}) = s(A_{n-k}) = A_k. \tag{19.262}$$

Donc O et A_k sont deux points fixes de l'isométrie f ; donc f est bien la réflexion sur le bon axe.

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

$$\frac{2(m-k)\pi}{n}, \tag{19.263}$$

et donc $f = r^{m-k}$, qui est de la forme demandée.

(2) Supposons à présent que f soit une réflexion d'axe Δ. Cette fois, Δ ne passe par aucun des points A_k, par contre Δ passe par 0. Nous commençons par montrer que Δ doit être la médiatrice d'un des côtés $[A_p, A_{p+1}]$ du polygone. Vu que Δ 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 Δ coupe $[A_p, A_{p+1}]$ en son milieu. Dans ce cas, Δ 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 Δ, nous avons aussi $d\big(f(A_p), \Delta\big) = \delta$ et $d\big(A_p, f(A_p)\big) = 2\delta$. D'autre part, par la définition de la distance, $\delta < x$. Si $x < \frac{l}{2}$, alors $\delta < \frac{\delta}{2}$ et donc $d\big(A_p, f(A_p)\big) < l$. 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

$$\det(sr^{n-2p-1}) = \det(s)\det(r^{n-2p-1}) = -1 \tag{19.264}$$

parce que $\det(s) = -1$ alors que $\det(r^k) = 1$ parce que r est une rotation dans $\mathrm{SO}(2)$. Ensuite nous avons

$$s \circ r^{n-2p-1}(A_p) = s(A_{p+n-2p-1}) = s(A_{n-p-1}) = A_{n-(n-p-1)} = A_{p+1}. \tag{19.265}$$

Donc $s \circ r^{n-2p-1}$ est bien une réflexion qui envoie A_p sur A_{p+1}.

□

Corollaire 19.95.
La liste des éléments de D_n est

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \tag{19.266}$$

et $|D_n| = 2n$.

Démonstration. Nous savons par la proposition 19.94 que tous les élément de D_n s'écrivent sous la forme r^k ou sr^k. Vu que r est d'ordre n, il ne faut considérer que $k \in \{1, \dots, n-1\}$. Les éléments 1, r,… , 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 :

$$sr^k(A_p) = s(A_{p+k}) = A_{n-p+k} \tag{19.267}$$

donc si $k \neq k'$, $sr^k(A_p) \neq sr^{k'}(A_p)$. La liste des éléments de D_n est donc

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \tag{19.268}$$

et donc $|D_n| = 2n$. □

Exemple 19.96
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 19.8. La symétrie d'axe vertical est nommée s et la rotation de 90 degrés est notée r.

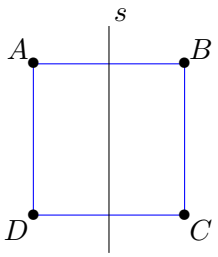


FIGURE 19.8 – Le carré dont nous étudions le groupe diédral.

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

$$(srs^{-1})(A, B, C, D) = sr(B, A, D, C) = s(A, D, C, B) = (B, C, D, A), \tag{19.269}$$

c'est-à-dire $srs^{-1} = r^{-1}$. Nous sommes alors dans le cadre du corollaire 2.78 et nous pouvons écrire que

$$D_4 = \mathrm{gr}(r) \times_\sigma \mathrm{gr}(s). \tag{19.270}$$

△

19.9.1.2 Table de multiplication

La proposition 19.94 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 $\mathbb{C}$.

Lemme 19.97.
Si R est une rotation autour de 0 (dans $\mathbb{C}$), et si s est la conjugaison complexe, alors

$$rs = sr^{-1} \tag{19.271}$$

Démonstration. Il s'agit seulement d'un calcul en écrivant R comme la multiplication par $e^{i\alpha}$. Nous avons

$$(Rs)z = e^{i\alpha}\bar{z} = s\big(e^{-i\alpha}z\big) = sR^{-1}z. \tag{19.272}$$

□

Proposition 19.98.
Si $\epsilon_1, \epsilon_2 \in \{0,1\}$ et si $k, l \in \mathbb{Z}$ nous avons

$$(s^{\epsilon_1}r^k)(s^{\epsilon_2}r^l) = s^{\epsilon_1+\epsilon_2}r^{l+(-1)^{\epsilon_1}k}. \tag{19.273}$$

Démonstration. Si $\epsilon_2 = 1$ alors nous utilisons le lemme 19.97 pour trouver

$$(s^{\epsilon_1}r^k)(s^{\epsilon_2}r^l) = s^{\epsilon_1}(r^k s^{\epsilon_2})r^l = s^{\epsilon_1}s^{\epsilon_2}r^{-k}r^l. \tag{19.274}$$

La proposition est déjà prouvée dans ce cas.

Passons à $\epsilon_2 = 0$. Dans ce cas nous avons

$$(s^{\epsilon_1}r^k)(s^{\epsilon_2}r^l) = s^{\epsilon_1}r^{k+l}, \tag{19.275}$$

et c'est bon. □

19.9.1.3 Générateurs : vue abstraite

19.99.
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 $n \geqslant 3$,
(3) $abab = e$.

Nous allons prouver que ce groupe doit avoir la même liste d'éléments que celle du corollaire 19.95.

Proposition 19.100 ([320])**.**
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 :

$$e \neq abab^{-1} = (ab)(ba^{-1})^{-1} = (ab)(ba)^{-1} \tag{19.276}$$

parce que $a^{-1} = a$. Donc $ab \neq ba$. □

Lemme 19.101 ([320])**.**
Pour tout k entre 1 et $n-1$ nous avons

$$\boldsymbol{Ad}(a)b^k = ab^k a^{-1} = ab^k a = b^{-k}. \tag{19.277}$$

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 $k+1$:

$$ab^{k+1}ba = ab^k ba = \underbrace{ab^k a}_{b^{-k}}\underbrace{aba}_{b^{-1}} = b^{-k}b^{-1} = b^{-(k+1)}. \tag{19.278}$$

□

Proposition 19.102.
L'élément a n'est pas une puissance de b.

Démonstration. Supposons le contraire : $a = b^k$. Dans ce cas nous aurions

$$e = (ab)(ab) = b^{k+1}b^{k+1} = b^{2k+2} = b^{2k}b^2 = a^2b^2 = b^2, \tag{19.279}$$

ce qui signifierait que b est d'ordre 2, ce qui est exclu par construction. □

Proposition 19.103 ([320]).
La liste des éléments de G est donnée par

$$G = \{1, b, \cdots, b^{n-1}, a, ab, \ldots, ab^{n-1}\} = \{a^\epsilon b^k\}_{\substack{\epsilon=0,1\\k=0,\ldots,n-1}} \tag{19.280}$$

Les éléments de ces listes sont distincts.

Démonstration. Étant donné que a n'est pas une puissance de b, les éléments 1, a, b,..., b^{n-1} sont distincts. De plus si k et $m = k + p$ sont deux éléments distincts de $\{1, \ldots, n-1\}$, 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 $1, a, b, \ldots, b^{n-1}, ab, \ldots, ab^{n-1}$ 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 $x = a^{m_1}b^{k_1}\ldots a^{m_r}b^{k_r}$ pour un certain r et avec pour tout i, $k_i \in \{1, \ldots, n-1\}$ (sauf k_r qui peut être égal à zéro) et $m_i = 1$, sauf m_1 qui peut être égal à zéro. Donc

$$x = a^m b^{k_1} ab^{k_2} a \ldots b^{k_{r-1}} ab^{k_r} \tag{19.281}$$

où m et k_r peuvent éventuellement être zéro. En utilisant le lemme 19.101 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. □

Lemme 19.104 ([1]).
Tout élément de G s'écrit de façon unique sous la forme $a^\epsilon b^k$ ou $b^k a^\epsilon$ avec $\epsilon = 0, 1$ et $k = 0, \ldots, n-1$.

Démonstration. Nous commençons par la forme $a^\epsilon b^k$. L'existence est la proposition 19.103. Pour l'unicité nous supposons $a^\epsilon b^k = a^\sigma b^l$ et nous décomposons en 4.

$\epsilon = 0, \sigma = 0$ Alors $b^k = b^l$. Mais b étant d'ordre n et k, l étant égaux au maximum à $n-1$, cette égalité implique $k = l$.

$\epsilon = 0, \sigma = 1$ Alors $b^k = ab^l$, ce qui donne $a = b^{k-l}$, ce qui est interdit par la proposition 19.102.

$\epsilon = 1, \sigma = 0$ Même problème.

$\epsilon = 1, \sigma = 1$ 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 $g = a^\epsilon b^k$ et 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 19.101

$$ab^k = b^{-k}a^{-1} = b^{-k}b^n a = b^{-k}a. \tag{19.282}$$

En ce qui concerne l'unicité, nous refaisons 4 cas pour $b^k a^\epsilon = b^l a^\sigma$ comme précédemment et ils se traitement exactement comme précédemment. □

Théorème 19.105.
Les groupes G et D_n sont isomorphes.

Démonstration. Nous utilisons l'application

$$\begin{aligned}\psi\colon G &\to D_n\\ a^k b^m &\mapsto s^k r^m.\end{aligned} \tag{19.283}$$

C'est évidemment bien défini et bijectif, mais c'est également un homomorphisme parce que si nous calculons ψ sur un produit, nous devons comparer

$$\psi\big(a^{k_1}b^{m_1}a^{k_2}b^{m_2}\big) \tag{19.284}$$

avec

$$\psi\big(a^{k_1}b^{m_1}\big)\psi\big(a^{k_2}b^{m_2}\big) = s^{k_1}r^{m_1}s^{k_2}r^{m_2}. \tag{19.285}$$

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 ψ dans (19.284) se simplifie en a^kb^m avec les même k et n que l'expression à droite dans (19.285) ne se simplifie en s^kr^m. □

Corollaire 19.106.
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 $srsr = e$.*
(2) *Le groupe D_n n'est pas abélien.*
(3) *Pour tout $k \in \{1, \dots, n-1\}$ nous avons $sr^ks = r^{-k}$.*
(4) *L'élément s ne peut pas être obtenu comme une puissance de r.*
(5) *La liste des éléments de D_n est*

$$D_n = \{1, r, \dots, r^{n-1}, s, sr, \dots, sr^{n-1}\} \tag{19.286}$$

(6) *Le groupe diédral D_n est d'ordre $2n$.*

Proposition 19.107.
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

$$D_n = C_n \times_\rho C_2 \tag{19.287}$$

où ρ désigne l'action adjointe.

Démonstration. L'isomorphisme est :

$$\begin{aligned}\psi\colon C_n \times_\rho C_2 &\to D_n\\ (b^k, a^\epsilon) &\mapsto b^ka^\epsilon.\end{aligned} \tag{19.288}$$

Action adjointe L'application $\rho_{a^\epsilon} = \mathbf{Ad}(a^\epsilon)$ est toujours un homomorphisme. Vu que a^ϵ est soit e soit a, nous allons nous restreindre à a et oublier l'exposant ϵ. Il faut montrer que $\mathbf{Ad}(a) \in \mathrm{Aut}(C_n)$. En utilisant le lemme 19.101,

$$\mathbf{Ad}(a)b^k = ab^ka^{-1} = b^{-k} = b^{n-k}. \tag{19.289}$$

L'application $\mathbf{Ad}(a)\colon C_n \to C_n$ est donc bijective et homomorphique. Ergo isomorphisme.

Injectif Si $\psi(b^k, a^\epsilon) = \psi(b^l, a^\sigma)$, alors par unicité du lemme 19.104 nous avons $k = l$ et $\epsilon = \sigma$.

Surjectif Par la partie « existence » du lemme 19.104.

Homomorphisme Lorqu'on prend deux sous-groupes d'un même groupe (ici le groupe des isométrique de $\mathbb{R}^2$), et que l'on tente de faire un produit demi-direct en utilisant l'action adjointe, nous avons toujours un homomorphisme. Dans notre cas, le calcul est :

$$\psi\big((b^k, a^\epsilon)(b^l, a^\sigma)\big) = b^k\rho_{a^\epsilon}(b^l)a^{\epsilon+\sigma} = b^ka^\epsilon b^la^{-\epsilon}a^{\epsilon+\sigma} = b^ka^\epsilon b^la^\sigma = \psi(b^k, a^\epsilon)\psi(b^l, a^\sigma). \tag{19.290}$$

□

19.9.1.4 Classes de conjugaison

Pour les classes de conjugaison du groupe diédral nous suivons [321].

D'abord pour des raisons de déterminants [18], 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

$$sr^k s = r^{-k} \tag{19.291a}$$

$$rs = sr^{-1} = sr^{n-1}. \tag{19.291b}$$

La classe de conjugaison qui ne rate jamais est bien entendu $C(1) = 1$. Nous commençons les vraies festivités $C(r^m)$. D'abord $r^k \cdot r^m = r^m$, ensuite

$$(sr^k) \cdot r^m = sr^k r^m r^{-k} s^{-1} = sr^m s^{-1} = r^{-m}. \tag{19.292}$$

Donc

$$C(r^m) = \{r^m, r^{-m}\}. \tag{19.293}$$

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

$$r^k \cdot s = r^k s r^{-k} = s s r^k s r^{-k} = s r^{-k} r^{-k} = s r^{n-2k}, \tag{19.294}$$

et

$$(sr^k) \cdot s = \underbrace{sr^k s}_{r^{-k}} r^{-k} s^{-1} = r^{-2k} s = r^{n-2k} s = s r^{(n-1)(n-2k)} = s r^{n^2-2kn-n+2k} = s r^{2k}. \tag{19.295}$$

donc

$$C(s) = \{sr^{n-2k}, sr^{2k}\}_{k=0,\dots,n-1}. \tag{19.296}$$

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

$$s \cdot (sr) = ssrs = rs = sr^{n-1}. \tag{19.297}$$

Ensuite

$$(sr^k) \cdot (sr) = sr^k s r r^{-k} s = r^{-2k+1} s = s r^{2k-1}. \tag{19.298}$$

Avec $k = \frac{n}{2}$, cela rend $s \cdot (sr)$, donc pas besoin de le recopier. Nous avons

$$C(sr) = \{sr^{2k-1}\}_{k=1,\dots,n-1}. \tag{19.299}$$

19.9.1.5 Le compte pour n pair

Si n est pair, nous avons les classes

$$C(1) = \{1\} \qquad \text{1 élément} \tag{19.300a}$$

$$C(r^m) = \{r^m, r^{m-1}\} \qquad \text{pour } 0 < m < \frac{n}{2} \qquad \frac{n}{2} - 1 \text{ fois 2 éléments} \tag{19.300b}$$

$$C(r^{n/2}) = \{r^{n/2}\} \qquad \text{1 élément} \tag{19.300c}$$

$$C(s) = \{sr^{2k}\}_{k=0,\dots,\frac{n}{2}-1} \qquad \frac{n}{2} \text{ éléments} \tag{19.300d}$$

$$C(sr) = \{sr^{2k+1}\}_{k=0,\dots,\frac{n}{2}-1} \qquad \frac{n}{2} \text{ éléments.} \tag{19.300e}$$

Au total nous avons bien listé $2n$ éléments comme il se doit, dans $\frac{n}{2} + 3$ classes différentes.

18. 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.

19.9.1.6 Le compte pour n impair

Si n est impair, nous avons les classes

$$C(1) = \{1\} \qquad 1 \text{ élément} \tag{19.301a}$$

$$C(r^m) = \{r^m, r^{m-1}\} \qquad \text{pour } 0 < m < \frac{n-1}{2} \qquad \frac{n-1}{2} \text{ fois 2 éléments} \tag{19.301b}$$

$$C(s) = \{sr^k\}_{k=0,\dots,n-1} \qquad n \text{ éléments} \tag{19.301c}$$

Au total nous avons bien listé $2n$ éléments comme il se doit, dans $\frac{n+3}{2}$ classes différentes.

19.9.2 Applications : du dénombrement

19.9.2.1 Le jeu de la roulette

Soit une roulette à n secteurs que nous voulons colorier en q couleurs[322]. Nous voulons savoir le nombre de possibilités à rotations près. Soit d'abord E l'ensemble des coloriages possibles sans contraintes ; 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 donné par le nombre d'orbites de l'action de G sur E qui sera donnée par la formule du théorème de Burnside 2.56.

Nous devons calculer $\text{Card}\big(\text{Fix}(g)\big)$ 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[19] $d$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 :

$$\frac{1}{n}\sum_{d|n}\varphi(d)q^{n/d}. \tag{19.302}$$

Cela est le nombre de coloriage possibles de la roulette à n secteurs avec q couleurs.

19.9.2.2 L'affaire du collier

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[20] D_n conservant un polygone a 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 par le milieu du côté opposé. Le groupe D_n contient n symétries axiales. Nous avons donc maintenant

$$|G| = 2n. \tag{19.303}$$

Nous écrivons la formule de Burnside

$$\text{Card}(\Omega) = \frac{1}{2n}\sum_{g\in G}\text{Card}\big(\text{Fix}(g)\big). \tag{19.304}$$

19. 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.

20. Définition 19.89.

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

$$qq^{(n-1)/2} = q^{\frac{n+1}{2}} \tag{19.305}$$

possibilités. Nous avons donc

$$\text{Card}(\Omega) = \frac{1}{2n}\left(\sum_{d|n} q^{n/d}\varphi(d) + nq^{\frac{n+1}{2}}\right). \tag{19.306}$$

Si n est pair, le choses se compliquent un tout petit peu. En plus de symétries axiales passant par un sommet et le milieu du côté opposé, il y a les axes passant par deux sommets opposés. Pour 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

$$q^2q^{\frac{n-2}{2}} = q^{\frac{n+2}{2}}. \tag{19.307}$$

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

$$\text{Card}(\Omega) = \frac{1}{2n}\left(\sum_{d|n} \varphi(d)q^{n/d} + \frac{n}{2}q^{(n+2)/2} + \frac{n}{2}q^{n/2}\right). \tag{19.308}$$

19.10 Classification des isométries dans $\mathbb{R}^2$

19.10.1 Réflexions

Soit un espace vectoriel E de dimension 2 muni d'un produit scalaire[21]. Cela pourrait très bien être $\mathbb{R}^2$, mais nous allons nous efforcer de l'appeler E pour rester un peu général.

Lemme-définition 19.108 (Caractérisation des réflexions)**.**
Soit une droite ℓ de $\mathbb{R}^2$. Il existe une unique application $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ telle que

(1) $f(x) = x$ pour tout $x \in \ell$.

(2) f échange les côtés de ℓ.

(3) f laisse invariants les droites perpendiculaires à ℓ et les cercles dont le centre est sur ℓ.

*Cette application est la **réflexion** d'axe ℓ.*

Démonstration. Soit x hors de ℓ et p la droite perpendiculaire à ℓ et passant par x. Nous avons $f(x) \in p$. En nommant P l'intersection entre ℓ et p, nous considérons le cercle $S(P, \|Px\|)$ qui est un cercle dont le centre est sur ℓ. 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 ℓ et f doit inverser les côtés de ℓ. Donc $f(x)$ est l'autre.

Cela prouve l'unicité. En ce qui concerne l'existence, il suffit de noter que la réflexion σ_ℓ satisfait les contraintes. □

Lemme 19.109.
Soit une droite ℓ est $A \in E$. Alors

$$\sigma_\ell(A) = 2\,\texttt{proj}_\ell(A) - A \tag{19.309}$$

où $\texttt{proj}_\ell$ est l'opération de projection orthogonale sur la droite ℓ.

21. Définition 11.5.

Démonstration. Nous posons $f(X) = 2\texttt{proj}_\ell(X) - X$ et nous allons montrer que $f = \sigma_\ell$ en vérifiant les conditions de la définition 19.108. Nous nous gardons bien de faire un raisonnement du type « nous allons montrer que f et σ_ℓ coïncident sur deux points, et sont donc égales par le corollaire 19.6 » parce que nous ne savons pas encore que σ_ℓ est une application affine, ni même que c'est une isométrie.

Si $X \in \ell$ alors $\texttt{proj}_\ell(X) = X$ et nous avons $f(X) = 2X - X = X$. Donc ℓ est conservée.

En ce qui concerne les deux côtés de ℓ, il existe une application linéaire $s\colon E \to \mathbb{R}$ et une constante $c \in \mathbb{R}$ telles qu'en posant $\ell(X) = s(X) + c$, la droite ℓ soit le lieux 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 :

$$\begin{aligned}
\ell\big(f(A)\big) &= \ell\big(2\texttt{proj}_\ell(A) - A\big) && \text{(19.310a)}\\
&= s(2\texttt{proj}_\ell(A) - A) + c && \text{(19.310b)}\\
&= 2s\big(\texttt{proj}_\ell(A)\big) - s(A) + c && \text{(19.310c)}\\
&= s\big(\texttt{proj}_\ell(A)\big) - s(A) && \text{(19.310d)}\\
&= -c - s(A) && \text{(19.310e)}\\
&= -\ell(A) && \text{(19.310f)}
\end{aligned}$$

où nous avons utilisé le fait que, $\texttt{proj}_\ell(A)$ étant sur ℓ, $s\big(\texttt{proj}_\ell(A)\big) + c = 0$. Nous avons donc $\ell\big(f(A)\big) = -\ell(A)$, ce qui indique que A et $f(A)$ sont de part et d'autre de ℓ.

Si d est une droite perpendiculaire à ℓ et si $A \in d$ alors $f(A) = 2\texttt{proj}_\ell(A) - A = \big(\texttt{proj}_\ell(A) - A\big) + A \in d$ parce que $\texttt{proj}_\ell(A) \in d$ du fait que d soit précisément perpendiculaire à ℓ. 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 $f(A) = S(K, r)$. Le segment $[A, f(A)]$ est par définition perpendiculaire à ℓ. Soit M, le milieu, qui est sur la droite ℓ. 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 donc préservé par f.

Nous en déduisons que $f = \sigma_\ell$. □

Proposition 19.110 ([323])**.**
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

$$B - A = \underbrace{B - \texttt{proj}_\ell(B)}_{=a} + \underbrace{\texttt{proj}_\ell(B) - \texttt{proj}_\ell(A)}_{=b} + \underbrace{\texttt{proj}_\ell(A) - A}_{=c}. \tag{19.311}$$

Vu que $b \perp a$ et $b \perp c$ nous avons

$$\|B - A\| = \langle B - A, B - A\rangle = \|a\|^2 + 2\langle a, c\rangle + \|b\|^2 + \|c\|^2. \tag{19.312}$$

Nous pouvons faire le même jeu avec $\sigma_\ell(B) - \sigma_\ell(A)$ en tenant compte du fait que $\texttt{proj}_\ell\big(\sigma_\ell(X)\big) = \texttt{proj}_\ell(X)$ et que

$$\sigma_\ell(A) - \texttt{proj}_\ell(A) = 2\texttt{proj}_\ell(A) - A - \texttt{proj}_\ell(A) = -\big(A - \texttt{proj}_\ell(A)\big). \tag{19.313}$$

Là nous avons utilisé le lemme 19.109. Ce que nous trouvons est que

$$\sigma_\ell(B) - \sigma_\ell(A) = -a + b - c, \tag{19.314}$$

et donc encore une fois

$$\|\sigma_\ell(B) - \sigma_\ell(A)\| = \|a\|^2 - 2\langle a, c\rangle + \|b\|^2 + \|c\|^2. \tag{19.315}$$

□

Remarque 19.111.
Il faut bien comprendre que si l'axe de la réflexion ne passe par par 0 (le zéro de l'espace vectoriel normé $(E, \|.\|)$), la réflexion n'est pas une isométrie de $(E, \|.\|)$ au sens où nous n'avons pas $\|\sigma_\ell(x)\| = \|x\|$.

Lemme 19.112.
Si A' est l'image de A par σ_ℓ alors ℓ est la médiatrice du segment $[A, A']$.

Démonstration. Soit $M \in \ell$. Nous avons

$$\|A - M\|^2 = \|\mathtt{proj}_\ell(A) - A\|^2 + \|\mathtt{proj}_\ell(A) - M\|^2 \tag{19.316}$$

parce que $A - \mathtt{proj}_{ell}(A) \perp M - \mathtt{proj}_\ell(A)$. Par ailleurs, vu que $\sigma_\ell(A) = 2\,\mathtt{proj}_\ell(A) - A$ et que $\mathtt{proj}_\ell(A) = \mathtt{proj}_\ell(A')$,

$$\|\mathtt{proj}_\ell(A) - A\| = \|\mathtt{proj}_\ell(A') - A'\|. \tag{19.317}$$

Nous avons donc

$$\|\sigma_\ell(A) - M\|^2 = \|A - M\|^2, \tag{19.318}$$

ce qui prouve que M est sur la médiatrice de $[A', A]$ par le lemme 19.9. □

19.113.
Si l est une droite dans $\mathbb{R}^2$, nous avons la réflexion $\sigma_l \in \mathrm{Isom}(\mathbb{R}^2)$ d'axe l. Cela est une isométrie et donc une application affine par le théorème 11.269. Le lemme suivant détermine comment la réflexion σ_ℓ se décompose en une translation et une application linéaire.

Lemme 19.114.
Soit une droite ℓ. Alors

$$\sigma_\ell = \tau_{2w} \circ \sigma_{\ell_0} \tag{19.319}$$

où ℓ_0 est la droite parallèle à ℓ passant par l'origine, et w est le vecteur perpendiculaire à ℓ tel que $\ell_0 = \ell + w$.

Démonstration. Il faut trouver trois points non alignés sur lesquels les deux applications coïncident ; cela suffira par le corollaire 19.6.

Pour tous les points de ℓ_0, l'égalité fonctionne parce que si $x \in \ell_0$,

$$\sigma_\ell(x) = x + 2w, \tag{19.320}$$

tandis que

$$\sigma_{\ell_0}(x) + 2w = x + 2w \tag{19.321}$$

du fait que $\sigma_{\ell_0}(x) = x$.

Si $x \in \ell$, alors

$$\sigma_\ell(x) = x \tag{19.322}$$

tandis que

$$\sigma_{\ell_0}(x) + 2w = x - 2w + 2w = x. \tag{19.323}$$

Donc les applications affines σ_ℓ et $x \mapsto \sigma_{\ell_0}(x) + 2w$ coïncident sur ℓ et ℓ_0. Elles coïncident donc partout. □

19.10.2 Translations et réflexions

Lemme 19.115.
Si $A \in \mathbb{R}^2$, si ℓ est une droite de $\mathbb{R}^2$, alors nous avons

$$\mathtt{proj}_{\tau_A(\ell)} = \tau_A \circ \mathtt{proj}_\ell \circ \tau_A^{-1}. \tag{19.324}$$

Démonstration. Soit $x \in \mathbb{R}^2$. Soit v unitaire dans la direction [22] de ℓ. La condition $q = \texttt{proj}_\ell(x)$ est le système

$$\begin{cases} q \in \ell & (19.325a) \\ (q-x) \cdot v = 0. & (19.325b) \end{cases}$$

Nous voulons prouver que $\tau_A(q) = \texttt{proj}_{\tau_A(\ell)}\big(\tau_A(x)\big)$, c'est à dire que

$$\begin{cases} \tau_A(q) \in \tau_A(\ell) & (19.326a) \\ \big(\tau_A(q) - \tau_A(x)\big) \cdot v = 0. & (19.326b) \end{cases}$$

Nous avons utilisé le fait que v est un vecteur unitaire dans la direction de $\tau_A(\ell)$ aussi bien que de ℓ.

Vu que $q \in \ell$, bien entendu que $\tau_A(q) \in \tau_A(\ell)$. D'autre part, $\tau_A(q) - \tau_A(x) = q - x$, donc

$$\big(\tau_A(q) - \tau_A(x)\big) \cdot v = (q-x) \cdot v = 0. \tag{19.327}$$

□

Lemme 19.116.
Si ℓ est une droite de $\mathbb{R}^2$, si $A \in \mathbb{R}^2$, alors

$$\sigma_{\tau_A(\ell)} = \tau_A \sigma_\ell \tau_A^{-1}. \tag{19.328}$$

Démonstration. Il s'agit d'un calcul mettant en scène les lemmes 19.109 et 19.115 :

$$\begin{aligned} \big(\sigma_{\tau_A(\ell)}\tau_A\big)(x) &= 2\,\texttt{proj}_{\tau_A(\ell)}\big(\tau_A(x)\big) & (19.329a) \\ &= 2\tau_A\big(\texttt{proj}_\ell(x)\big) - \tau_A(x) & (19.329b) \\ &= 2\,\texttt{proj}_\ell(x) + 2A - x - A & (19.329c) \\ &= 2\,\texttt{proj}_\ell(x) - x + A & (19.329d) \\ &= \tau_A\big(\sigma_\ell(x)\big). & (19.329e) \end{aligned}$$

□

19.10.3 Rotations

Définition 19.117 (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.*

19.118.
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 19.129 nous dirons le plus souvent « rotation » pour « rotation centrée en 0 ». D'où les énoncés comme « les rotations sont les matrice orthogonales » (corollaire 19.137), qui *stricto senus* de la définition 19.117 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.

22. 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 contentent du cas vectoriel, de nombreuses choses sont plus simples.

Corollaire 19.119.
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 19.112 nous dit que si A est envoyé sur B, l'axe est forcément la médiatrice de $[A, B]$. □

Lemme-définition 19.120 ([170])**.**
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 $\sigma_1 \circ r = s_2$ et donc

$$\sigma_1(A) = (\sigma_1 \circ r)(A) = s_2(A). \tag{19.330}$$

On pose $B = \sigma_1(A)$. Alors σ_1 et σ_2 envoient tout deux A sur B.

Si $A = B$ alors A est fixé par σ_1 et donc appartient à ℓ_1. Même chose pour A est fixé par σ_2 et donc $A \in \ell_2$. Cela donne $A = B = O$, et donc le point fixé par r est O.

Si $A \neq B$ alors il existe une unique réflexion envoyant A sur B (corollaire 19.119). L'unicité signifie que $\sigma_1 = \sigma_2$. Dans ce cas, $r = \sigma_1 \circ \sigma_2 = \mathrm{Id}$. □

19.121.
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 σ_1 et σ_2 d'être identiques à ceux que σ_a et σ_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 les définition suivante.

Lemme 19.122.
Les rotation 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 σ_1 et σ_2 sont des isométries de (E, d) (19.110) nous avons :

$$d(X, Y) = d\big(\sigma_2(X), \sigma_2(Y)\big) = d\big(\sigma_1\sigma_2(X), \sigma_1\sigma_2(Y)\big) = d\big(r(X), r(Y)\big). \tag{19.331}$$

□

Ce lemme nous dit qu'une rotation de centre O vérifie $\|OX\| = \|Or(X)\|$ pour tout X.

Proposition 19.123 ([170])**.**
Soient $A, B, O \in E$ tels que $\|AO\| = \|BO\| \neq 0$. Alors il existe une unique rotation r centrée en O telle que $r(A) = B$.

Problèmes et choses à faire

Attention : la preuve qui suit contient de nombreuses galipettes et improvisations personnelles. Relisez-la attentivement avant de la prendre pour argent comptant.

La difficulté tient essentiellement à ce que cette preuve traite de façon vectorielle (tous les points sont des éléments de E) un énoncé qui est essentiellement affine : tous les points doivent être vus comme vecteurs partant de O.

Si vous comparez la preuve donnée ici avec celle de [170], vous remarquerez que dans ce dernier, seule la partie « A et O son alignés » est présente. C'est parce que lui, il se met directement dans le cas vectoriel et $O = 0$. Il a donc une preuve un tout petit peu moins générale, mais au moins ses isométries sont linéaires et non affines.

Démonstration. Existence et unicité séparément.

Existence Si $A = B$, l'identité fait l'affaire. Sinon, $\|A-O\| = \|B-O\|$ implique que la médiatrice de $[A, B]$ contient O. Soit σ_m la réflexion selon cette médiatrice. La rotation $\sigma_m \circ \sigma_{(AO)}$ convient.

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 19.120. 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.

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 $g = \sigma_{(AO)}$.

Si A et O sont alignés Soit maintenant un point C tel que $A - O \perp C - O$ et

$$\|OC\| = \|OA\| = \|OB\|. \tag{19.332}$$

Vu que 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 $(E, \|.\|)$:

$$\|X - Y\| = \|g(X) - g(Y)\| = \|s(X) + G - s(y) - G\| = \|s(X) - s(Y)\| = \|s(X - Y)\| \tag{19.333}$$

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

$$O - s(O) = A - s(A). \tag{19.334}$$

Vu que s est une isométrie (une vraie) nous avons

$$s(A - O) \perp s(C - O), \tag{19.335}$$

mais $s(A - O) = s(A) - s(O) = A - O$ par (19.334). Donc

$$A - O \perp s(C - O). \tag{19.336}$$

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

$$g(A - O) = s(A) - s(O) + G = s(A) - O + O - s(O) + G = A - O + G \tag{19.337a}$$
$$g(C - O) = C - O + G, \tag{19.337b}$$

ce qui signifierait que g et τ_G coïncideraient sur les points $A-O$ et $C-O$, et donc seraient égaux par le corollaire 19.6. 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

$$\sigma_{(AO)}(A-O)=A-O \tag{19.338}$$

et parce que $C-O$ est perpendiculaire à cette droite :

$$\sigma_{(AO)}(C-O)=-(C-O). \tag{19.339}$$

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

$$\sigma_{(OA)}=\sigma_m\circ r, \tag{19.340}$$

et donc r est fixé à

$$r=\sigma_m\circ\sigma_{(OA)}. \tag{19.341}$$

□

19.124.
Anticipons un peu et faisons semblant de déjà connaitre les matrices et les fonctions trigonométriques. La proposition 19.123 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 $\mathbb{R}^2\to\mathbb{R}^2$, 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.

Proposition 19.125 ([170]).
Soit une rotation r de E centrée en O. Pour toute réflexion σ_ℓ telle que le centre de r soit sur ℓ, il existe une réflexion σ_1 tells que $r=\sigma_1\circ\sigma_\ell$. Il existe aussi une réflexion σ_2 telle que $r=\sigma_\ell\circ s_2$.

Démonstration. Si $r=\mathrm{Id}$ c'est bon avec $s_1=s_2=\sigma_\ell$. Sinon nous considérons $A\neq O$ sur ℓ, et $B=r(A)$. Nous savons que $B\neq A$ parce que O est le seul point de E fixé par r (proposition 19.120). Il existe une réflexion (unique) σ_1 faisant $\sigma_1(A)=B$, et c'est le 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 19.122).

La rotation $\sigma_1\circ\sigma_\ell$ vérifie

$$(\sigma_1\circ\sigma_\ell)(A)=\sigma_1(A)=B. \tag{19.342}$$

Or $\|OA\|=\|OB\|$, donc il y a unicité de la rotation centrée en O portant A sur B (proposition 19.123) ; 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 σ_2 tel que $r^{-1}=\sigma_2\circ\sigma_\ell$, ce qui donne

$$r=\sigma_\ell\circ\sigma_2. \tag{19.343}$$

□

19.10.4 Rotation d'un angle donné

Lemme-définition 19.126.
Soit $\theta \in \mathbb{R}$. Nous considérons l'application linéaire $R_0(\theta)\colon \mathbb{R}^2 \to \mathbb{R}^2$ dont la matrice dans la base canonique est

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{19.344}$$

(1) Nous avons

$$R_0(\theta) = \sigma_\ell \circ s \tag{19.345}$$

où s est la réflexion d'axe horizontal et ℓ est la droite

$$\ell = \mathbb{R} \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix}. \tag{19.346}$$

(2) L'application $R_0(\theta)$ est une rotation autour de $(0,0)$.

(3) Si $A \in \mathbb{R}^2$, alors l'application

$$R_A(\theta) = \tau_A \circ R_0(\theta) \circ \tau_A^{-1} \tag{19.347}$$

est une rotation autour de A nommée ***rotation d'angle*** *θ.*

Démonstration. Nous allons prouver l'égalité (19.345) en calculant les deux membres sur les vecteurs $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ qui forment une base.

Pour $p = (1,0)$ D'abord la chose facile[23] : $s(1,0) = (1,0)$.

Pour calculer $\sigma_\ell(1,0)$, nous utilisons le lemme 19.109 ; nous commençons par chercher la projection orthogonale q de $p = (1,0)$ sur ℓ. Nous posons

$$q = \lambda \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \tag{19.348}$$

et nous cherchons λ satisfaisant $q \cdot (q-p) = 0$. Un peu de calculs passant par (19.58) nous donne

$$q \cdot (q-p) = \lambda\big(\lambda - \cos(\theta/2)\big). \tag{19.349}$$

Les deux solutions sont $\lambda = 0$ et $\lambda = \cos(\theta/2)$. Mais la solution $\lambda = 0$ revient à dire que la droite ℓ est verticale, c'est à dire $\cos(\theta/2) = 0$. Donc la solution est toujours donnée par

$$\lambda = \cos(\theta/2). \tag{19.350}$$

Nous introduisons cette valeur dans (19.348) pour fixer q, et nous utilisons la formule du lemme 19.109 :

$$\sigma_\ell \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2q - p = \begin{pmatrix} 2\cos^2(\theta/2) - 1 \\ 2\cos(\theta/2)\sin(\theta/2) \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}. \tag{19.351}$$

Nous avons utilisé le formules trigonométrique de duplication d'angle (corollaire 19.27).

Pour $p = (0,1)$ Cette fois $s(p) = (0,-1)$ et l'équation pour déterminer λ est

$$0 = q \cdot (q-p) = \begin{pmatrix} \lambda\cos(\theta/2) \\ \lambda\sin(\theta/2) \end{pmatrix} \cdot \begin{pmatrix} \lambda\cos(\theta/2) \\ \lambda\sin(\theta/2) - 1 \end{pmatrix}. \tag{19.352}$$

23. 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 du 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 trouvons $\lambda\big(\lambda+\sin(\theta/2)\big)=0$. Le cas $\lambda=0$ signifie que la droite ℓ est horizontale et donc que $\sin(\theta/2)=0$. Donc la solution est dans tous les cas

$$\lambda=-\sin(\theta/2). \tag{19.353}$$

Nous trouvons

$$\sigma_\ell\begin{pmatrix}0\\-1\end{pmatrix}=2q-p=\begin{pmatrix}-\sin(\theta)\\\cos(\theta)\end{pmatrix}. \tag{19.354}$$

Le calcul de $\sigma_\ell\circ s$ étant fait sur une base, il est facile de reconstituer la matrice

$$\begin{pmatrix}\cos(\theta)&-\sin(\theta)\\\sin(\theta)&\cos(\theta)\end{pmatrix}. \tag{19.355}$$

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).

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 ℓ' 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

$$R_A(\theta)=\tau_A\sigma_\ell\sigma_{\ell'}\tau_A^{-1}=\underbrace{\tau_A\sigma_\ell\tau_A^{-1}}_{=\sigma_{\tau_A(\ell)}}\underbrace{\tau_A\sigma_{\ell'}\tau_A^{-1}}_{\sigma_{\tau_A(\ell')}}=\sigma_{\tau_A(\ell)}\sigma_{\tau_A(\ell')}. \tag{19.356}$$

Nous avons utilisé le lemme 19.116.

Nous voyons que $R_A(\theta)$ est une composée de deux réflexions se coupant en A. C'est donc une rotation centrée en A.

□

19.10.5 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 19.127.
Quelque résultats à propos de rotations.

(1) *Toutes les rotations de $\mathbb{R}^2$ centrées en 0 sont de la forme $R_0(\theta)$ pour un $\theta\in\mathbb{R}$.*

(2) *Les rotations de $\mathbb{R}^2$ centrées en 0 sont des applications linéaires.*

(3) *Si r est une rotation dans $\mathbb{R}^2$, il existe $A\in\mathbb{R}^2$ et $\theta\in\mathbb{R}$*

$$r=\tau_A\circ R_0(\theta)\circ\tau_A^{-1}. \tag{19.357}$$

Démonstration. Soit r une rotation centrée en $(0,0)$. La proposition 19.125, il existe une droite ℓ passant par $(0,0)$ telle que telle que $r=\sigma_\ell\circ s$ où s est la réflexion d'axe horizontal : $s(x,y)=(x,-y)$.

Soit un vecteur unitaire $v\in\ell$. Vu que $v\in S^1$, la proposition 19.58 nous donne $t\in[0,2\pi[$ tel que

$$v=\begin{pmatrix}\cos(t)\\\sin(t)\end{pmatrix}. \tag{19.358}$$

En posant $\theta=2t$ nous avons

$$\ell=\mathbb{R}\begin{pmatrix}\cos(\theta/2)\\\sin(\theta/2)\end{pmatrix}, \tag{19.359}$$

et donc $r=R_0(\theta)$ qui est l'application linéaire définie en 19.126.

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 ℓ_1 et ℓ_2 sont des droites telles que $\ell_1\cap\ell_2=\{A\}$.

En utilisant le lemme 19.116, nous avons $\sigma_{\ell_i} = \tau_A \circ \sigma_{\tau_A^{-1}(\ell_i)} \circ \tau_A^{-1}$. En substituant cela et en nous rendant compte que $\tau_A^{-1}\tau_A = \text{Id}$ nous avons

$$r = \sigma_{\ell_1} \circ \sigma_{\ell_2} = \tau_A r_0 \tau_1^{-1} \tag{19.360}$$

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 θ.

□

Proposition 19.128 ([170]).
Les rotations basées en O 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.

Composition Soient des rotations r, r' centrées en O. Soit également une réflexion σ dont l'axe contient O. Alors la proposition 19.125 nous donne l'existence de σ_1 et σ_2 tels que $r = \sigma_1\sigma$ et $r' = \sigma\sigma_2$. Avec ça, la composition donne

$$rr' = \sigma_1\sigma\sigma\sigma_2 = \sigma_1\sigma_2, \tag{19.361}$$

qui est encore une rotation.

Commutativité Soient deux rotations r et r' ainsi que des décompositions $r = \sigma_1\sigma$, $r' = \sigma\sigma_2$. Nous avons

$$rr' = \sigma_1\sigma_2 \tag{19.362a}$$
$$r'r = \sigma\sigma_2\sigma_2\sigma. \tag{19.362b}$$

Vu que $t = \sigma_2\sigma_1$ est une rotation nous pouvons encore appliquer la proposition 19.125 pour avoir $t = \sigma_2\sigma_1 = \sigma\sigma_3$. Avec ça,

$$r'r = \sigma\sigma\sigma_3\sigma = \sigma_3\sigma. \tag{19.363}$$

Mais aussi $rr' = \sigma_1\sigma_2 = t^{-1} = \sigma_3\sigma$. Nous avons donc bien $rr' = r'r$, et le groupe est commutatif.

□

19.129.
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 ».

Vu que nous allons maintenant prendre un point de vue plus vectoriel, nous allons noter les points de E avec des lettres comme x, y, u, v et plus avec des majuscules, comme quand on avait un point de vue afin. En même temps, nous allons noter les applications $E \to E$ par des lettres comme A et ne plus écrire les parenthèses. Bref, nous écrivons Au au lieu de $r(A)$.

Lemme 19.130.
En dimension 2, les réflexions vectorielles (c'est-à-dire dont l'axe passe par 0) ont un déterminant -1.

Démonstration. Soit une réflexion d'axe ℓ. Prenons une base orthonormale de E constituée de e_1 sur ℓ 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

$$\det(\sigma_\ell) = e_1^*\big(\sigma_\ell(e_1)\big)e_1^*\big(\sigma_\ell(e_2)\big) - e_2^*\big(\sigma_\ell(e_1)\big)e_1^*\big(\sigma_\ell(e_2)\big) = 1 \times (-1) - 0 \times 0 = -1. \tag{19.364}$$

Nous utilisons de façon cruciale le fait que le calcul du déterminant ne dépende pas de la base choisie, lemme 11.50. □

Proposition 19.131.
Les rotations[24] *sont*

(1) *des applications linéaires orthogonales au sens de la définition 11.80,*

(2) *des applications de déterminant* 1,

Démonstration. Le fait qu'elles soient linéaires est la proposition 19.127.

Nous avons, pour tout $u \in E$ l'égalité de la norme $\|u\|$ et $\|Au\|$ par le lemme 19.122 appliqué à $Y = 0$. En terme de produits scalaires nous avons alors $\langle Au, Au\rangle = \langle u, u\rangle$, et donc

$$\langle A^*Au, u\rangle = \|u\|^2. \tag{19.365}$$

En particulier si $\{e_i\}_{i=1,\dots,n}$ est une base orthonormée de E nous avons

$$(A^*Ae_i)_i = \|e_i\|^2 = 1, \tag{19.366}$$

ce qui donne $\|A^*Ae_i\| \geqslant 1$, avec égalité si et seulement si $A^*Ae_i = e_i$. Ici nous avons utilisé le fait que $\langle x, e_i\rangle = x_i$, et le fait que pour tout i nous ayons $\|x\| \geqslant |x_i|$, avec égalité seulement si x est un multiple de e_i.

Par ailleurs l'inégalité de Cauchy-Schwarz 11.10 nous donne

$$\|u\|^2 = |\langle A^*Au, u\rangle| \leqslant \|A^*Au\|\|u\| \tag{19.367}$$

et donc

$$\|u\| \leqslant \|A^*Au\|. \tag{19.368}$$

Encore une fois, en appliquant cela à $u = e_i$ nous trouvons $1 \leqslant \|A^*Ae_i\|$. Vu que nous avions déjà l'inégalité dans l'autre sens, $\|A^*Ae_i\| = 1$. Et le cas d'égalité est uniquement possible avec $A^*Ae_i = e_i$.

Donc pour tout i de la base nous avons $A^*Ae_i = e_i$. Nous avons donc $A^*A = \mathrm{Id}$ et l'application A est orthogonale.

En ce qui concerne le déterminant, les réflexions sont de déterminant -1 par le lemme 19.130, donc $A = \sigma_1 \circ \sigma_2$ est de déterminant 1. Nous avons utilisé le fait que le déterminant était un morphisme : proposition 11.52(1). □

Remarque 19.132.
Nous ne savons pas encore que les rotations forment tout le groupe $\mathrm{SO}(2)$ des endomorphismes orthogonaux de déterminant 1. Il faudra attendre le corollaire 19.137 pour le savoir.

Lemme 19.133.
L'application $-\mathrm{Id}$ *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 σ_i est la réflexion le long de l'axe $\ell_i = \{te_i\}_{t\in\mathbb{R}}$. Faut-il vous prouver que $r = -\mathrm{Id}$? La réflexion σ_2 retourne la composante y d'un vecteur écrit dans la base $\{e_1, e_2\}$ sans toucher à la composante x. La réflexion σ_1 fait le contraire. □

19.10.6 Matrice des transformations orthogonales

Nous donnons maintenant une forme générale (trigonométrique) pour les matrices de $\mathrm{SO}(2)$. Nous ne pouvons cependant pas invoquer les lemmes 19.134 ou 19.135 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 19.138.

Lemme 19.134.
Si $A \in \mathrm{O}(2)$ *alors il existe un unique* $\theta \in [0, 2\pi[$ *et un unique* $\epsilon = \pm 1$ *tels que*

$$A = \begin{pmatrix} \cos(\theta) & -\epsilon\sin(\theta) \\ \sin(\theta) & \epsilon\cos(\theta) \end{pmatrix} \tag{19.369}$$

24. Centrées en 0, nous ne le répéterons pas !

Démonstration. Soit une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et imposons qu'elle soit dans $\mathrm{O}(2)$. Le fait que A soit orthogonale impose

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{19.370}$$

Nous avons alors le système

$$\begin{cases} a^2 + b^2 = 1 & \text{(19.371a)} \\ b^2 + d^2 = 1 & \text{(19.371b)} \\ ab + cd = 0 & \text{(19.371c)} \end{cases}$$

La proposition 19.58 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 θ. Le α ainsi fixé ne sera peut-être pas dans $[0, 2\pi[$, mais ce ne sera pas grave.

Les angles θ et α sont alors liés par la contrainte

$$\cos(\theta)\cos(\alpha) + \sin(\theta)\sin(\alpha) = 0. \tag{19.372}$$

Utilisant l'identité (19.72c) cela signifie que $\cos(\theta - \alpha) = 0$. Donc

$$\alpha \in \{\theta + \frac{\pi}{2} + k\pi\}_{k\in\mathbb{Z}}. \tag{19.373}$$

Si k est pair, ça donne

$$\cos(\alpha) = -\sin(\theta) \tag{19.374a}$$
$$\sin(\alpha) = \cos(\theta) \tag{19.374b}$$

et alors

$$A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{19.375}$$

Si au contraire k est impair, alors

$$\cos(\alpha) = \sin(\theta) \tag{19.376a}$$
$$\sin(\alpha) = -\cos(\theta), \tag{19.376b}$$

et

$$A = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}. \tag{19.377}$$

Nous avons démontré qu'une matrice de $\mathrm{O}(2)$ était forcément d'une des deux formes (19.375) ou (19.377). Il est maintenant facile de vérifier que ces deux matrices sont effectivement dans $\mathrm{O}(2)$. □

Lemme 19.135.
Tout élément de $\mathrm{SO}(2)$ *s'écrit (dans la base canonique) de façon unique sous la forme*

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{19.378}$$

avec $\theta \in [0, 2\pi[$.

Démonstration. Vu que SO(2) est la partie de O(2) constitué des matrices de déterminant 1, nous pouvons reprendre la forme donnée par le lemme 19.134 et fixer ϵ par la contrainte sur le déterminant.

Nous avons, en utilisant la relation du lemme 19.18,

$$\det\begin{pmatrix}\cos(\theta) & -\epsilon\sin(\theta)\\ \sin(\theta) & \epsilon\cos(\theta)\end{pmatrix}=\epsilon, \tag{19.379}$$

et donc il faut et suffit de fixer $\epsilon=1$. □

Corollaire 19.136 ([1]).
Nous avons une bijection

$$\begin{aligned}\psi\colon \mathrm{SO}(2)&\to[0,2\pi[\\ \begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}&\mapsto\theta,\end{aligned} \tag{19.380}$$

et un isomorphisme de groupe

$$\begin{aligned}\varphi\colon \mathrm{SO}(2)&\to \mathrm{U}(1)=\{e^{i\theta}\}_{\theta\in\mathbb{R}}\\ \begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}&\mapsto e^{i\theta}.\end{aligned} \tag{19.381}$$

Démonstration. La première assertion est une paraphrase du lemme 19.135. 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 19.26 qui jouent. En ce qui concerne la bijection...

Surjection Vu que $e^{i\theta+2ki\pi}=e^{i\theta}$, tout élément de U(1) est exponentielle de $i\theta$ pour un $\theta\in[0,2\pi[$.

Injection Nous avons $\varphi(A)=\varphi(B)$ lorsque les formes matricielles de A et B sous forme trigonométrique sont avec des angles différents d'un multiple de 2π. Vu que les fonctions trigonométriques sont périodiques, nous avons $A=B$ parce que leurs matrices sont égales.

□

19.10.7 Rotations, SO(2) et matrice de rotation

Corollaire 19.137 ([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 19.127 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

$$\begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix}. \tag{19.382}$$

Donc toute rotation est dans SO(2).

D'autre part, le lemme 19.135 indique que tout élément de SO(2) a, dans la base canonique, une matrice de la forme (19.382). Le lemme 19.126 nous indique alors que c'est une rotation. □

Proposition 19.138.
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 corollaire 19.137 nous dit qu'il existe $\theta \in \mathbb{R}$ tel que

$$r = R_0(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{19.383}$$

Nous avons identifié l'application linéaire à sa matrice. L'élément $\big(\cos(\theta), \sin(\theta)\big)$ est dans S^1, et il existe donc, par la proposition 19.58, un unique $t \in [0, 2\pi[$ tel que $\cos(t) = \sin(\theta)$ et $\sin(t) = \sin(\theta)$. Pour ce t nous avons alors

$$r = R_0(\theta) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}. \tag{19.384}$$

□

Proposition 19.139.
Nous avons la formule suivante pour la composition :

$$R_0(\alpha) \circ R_0(\beta) = R_0(\alpha + \beta). \tag{19.385}$$

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 19.26 :

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} \tag{19.386a}$$

$$= \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\beta) \end{pmatrix} \tag{19.386b}$$

$$= \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}. \tag{19.386c}$$

Donc dans la base canonique, la matrice de $R_0(\alpha)R_0(\beta)$ est celle de $R_0(\alpha+\beta)$. □

19.10.8 Rotation et application affine

Nous considérons à nouveau la définition 10.8 d'une application affine ainsi que sa décomposition en application linéaire et translation donnée par le lemme 10.9. Nous voyons maintenant comment ces choses se mettent dans le cas d'une rotation non centrée en l'origine.

Exemple 19.140
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[25]. Nous allons extraire de f la partie linéaire définie en 10.9.

Il existe des axes ℓ_1 et ℓ_2 tels que $\ell_1 \cap \ell_2 = \{A\}$ et tels que

$$f = s_{\ell_1} \circ s_{\ell_2}. \tag{19.387}$$

En utilisant le lemme 19.116,

$$f = t_A \circ s_{t_A^{-1}(\ell_1)} \circ t_A^{-1} \circ t_A \circ s_{t_A^{-1}(\ell_2)} \circ t_A^{-1} = t_A \circ s_{t_A^{-1}(\ell_1)} \circ s_{t_A^{-1}(\ell_2)} \circ t_A^{-1}. \tag{19.388}$$

Vu que $A \in \ell_i$, 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

$$f = t_A \circ R \circ t_A^{-1} \tag{19.389}$$

où R est une rotation centrée en O ; donc une application linéaire par la proposition 19.127.

25. Voir la définition 19.117.

Nous avons

$$\begin{aligned} f(x) &= (t_A \circ R \circ t_A^{-1})(x) && (19.390a)\\ &= (t_A \circ R)(x - A) && (19.390b)\\ &= t_A\big(R(x) - R(A)\big) && (19.390c)\\ &= R(x) + t_{R(A)+A} && (19.390d)\\ &= (t_{R(A)+A} \circ R)(x). && (19.390e) \end{aligned}$$

Donc

$$f = t_{R(A)+A} \circ R. \tag{19.391}$$

Il est maintenant aisé de montrer que R est la partie linéaire de f. Pour tout $M, x \in \mathbb{R}^2$ nous avons

$$\begin{aligned} f(M + x) &= R(M + x) + R(A) + A && (19.392a)\\ &= R(M) + R(x) + R(A) + A && (19.392b)\\ &= R(x) + f(M). && (19.392c) \end{aligned}$$

Donc ok pour la formule

$$f(M + x) = R(x) + f(M) \tag{19.393}$$

et R est la partie linéaire de f, voir la définition 10.9. △

Lemme 19.141 ([1]).
Tout sous-groupe fini de SO(2) *est cyclique.*

Démonstration. Soit uns sous-groupe fini G de SO(2). Nous savons que SO(2) est isomorphe à U(1) par le corollaire 19.136, et en bijection avec $[0, 2\pi[$. Vu que G est fini, l'ensemble $G\backslash\{e\}$ il possède, dans $[0, 2\pi[$ un élément minimum non nul. Soit g_0 ce minimum.

Soit un élément g_1 de G qui ne serait ni l'identité ni un multiple de g_0. En particulier tous les nombres du type $g_1 - kg_0$ sont dans G (l'image de G dans $[0, 2\pi[$ en fait). Si g_1 n'est pas un multiple de g_0, il n'en reste pas moins que $g_1 = \lambda g_0$; alors en prenant pour k la partie entière de λ, l'élément $g_1 - kg_0$ est plus petit que g_0. Contradiction. □

19.10.9 Angle entre deux droites

Avant d'aborder la classification des isométries, nous devons parler de l'angle entre deux droites. Si ℓ_1 et ℓ_2 sont deux droites, alors il est bien clair deux angles peuvent prétendre être « l'angle entre ℓ_1 et ℓ_2 ». De plus chacun de ces deux angles sont doubles parce que si α peut prétendre être l'angle entre ℓ_1 et ℓ_2, alors $-\alpha$ peut également prétendre.

Remarque 19.142.
Nous ne parlons pas de l'angle entre ℓ_1 et ℓ_2 mais bien de l'angle *de* ℓ_1 *à* ℓ_2. L'ordre des droites est important.

19.143.
Pour la suite, $R_O(\alpha)$ est la rotation d'angle α autour du point O tandis que $R(\alpha)$ est la rotation d'angle α autour de l'origine.

Proposition 19.144 ([324]).
Si u et v sont des vecteurs unitaires[26] *de $\mathbb{R}^2$ alors il existe une unique rotation*[27] *f telle que $f(u) = v$.*

Démonstration. C'est la proposition 19.123 appliquée à $O = (0, 0)$. □

26. De norme 1.
27. Définition 19.117.

Remarque 19.145.
Notons l'unicité. Nous ne faisons pas de différences entre R_θ 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 θ de façon univoque. Par contre ça définit une classe modulo 2π, c'est-à-dire un élément $\theta \in \mathbb{R}/2\pi$.

Nous avons déjà défini le groupe $\mathrm{SO}(2)$ en la définition 11.84 et nous avons déterminé ses matrices dans $\mathbb{R}^2$ en le lemme 19.135.

19.10.10 Angle orienté

La proposition 19.144 donne une application

$$T\colon S^1 \times S^1 \to \mathrm{SO}(2). \tag{19.394}$$

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 $g \in \mathrm{SO}(2)$ telle que $g(u) = u'$ et $g(v) = v'$.

Définition 19.146 (Angle orienté[324])**.**
*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).*

Proposition 19.147.
Nous avons $T(u,v) = T(u',v')$ si et seulement si $(u,v) \sim (u',v')$.

Démonstration. En utilisant la commutativité du groupe $\mathrm{SO}(2)$ 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 19.148.
Nous avons une bijection

$$\begin{aligned} S\colon \frac{S^1 \times S^1}{\sim} &\to \mathrm{SO}(2) \\ [u,v] &\mapsto T(u,v). \end{aligned} \tag{19.395}$$

Démonstration. En plusieurs points.

S est bien définie En effet si $[u,v] = [z,t]$ alors $T(u,v) = T(z,t)$.

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 19.147.

Surjectif Nous avons $R_\theta = T(u, R_\theta u)$.

□

Définition 19.149 (Somme d'angles orientés[324])**.**
Si $[u,v]$ et $[z,t]$ sont des angles orientés, nous définissons la somme par

$$[u,v] + [z,t] = S^{-1}\Big(S[u,v] \circ S[z,t]\Big). \tag{19.396}$$

Lemme 19.150.
Quelques propriétés des angles plats liées à la somme.

(1) $(S^1 \times S^1)/\sim$ est un groupe commutatif.

(2) Relations de Chasles :

$$[u,v] + [v,w] = [u,w]. \tag{19.397}$$

(3) $-[u,v]=[v,u]$.

Démonstration. Pour la relation de Chasles, ça se base sur la propriété correspondante sur T :

$$\begin{aligned}[u,v]+[v,w]&=S^{-1}\Big(T(u,v)\circ T(v,w)\Big) &(19.398a)\\ &=S^{-1}\big(T(u,w)\big) &(19.398b)\\ &=[u,w]. &(19.398c)\end{aligned}$$

Pour l'inverse, la vérification est que

$$[u,v]+[v,u]=[u,u]=0. \tag{19.399}$$

□

Définition 19.151.
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 $\mathrm{SO}(2)$ ne définit pas θ, mais seulement sa classe modulo 2π. C'est pour cela que la mesure de l'angle orienté n'est également définie que modulo 2π.

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

$$\overrightarrow{AB}=\frac{B-A}{\|B-A\|}. \tag{19.400}$$

C'est le vecteur unitaire dans la direction « de B vers A ».

Théorème 19.152 (Théorème de l'angle inscrit[325])**.**
Soit un cercle Γ *de centre* O *et trois points distincts* $A,B,M\in\Gamma$. *Alors*

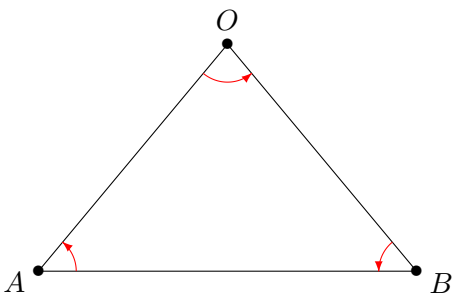
$$2(\overrightarrow{MA},\overrightarrow{MB})\in(\overrightarrow{OA},\overrightarrow{OB})_{2\pi} \tag{19.401}$$

où l'indice 2π *indique la classe modulo* 2π.

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 π, mais comme nous n'avons pas défini les angles autrement que modulo π, nous ne pouvons pas dire mieux. Donc

$$2(\overrightarrow{AB},\overrightarrow{AO})+(\overrightarrow{OB},\overrightarrow{OA})\in[\pi]_{2\pi}. \tag{19.402}$$

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

$$2(\overrightarrow{MB},\overrightarrow{MO})+(\overrightarrow{OM},\overrightarrow{OB})\in[\pi]_{2\pi}. \tag{19.403}$$

Nous faisons la différence entre les deux équations en remaquant que la différence de deux représentants de $[\pi]_{2\pi}$ est un représentant de $[0]_{2\pi}$ et en en nous souvenant que $-(\overrightarrow{MB},\overrightarrow{MO})=(\overrightarrow{MO},\overrightarrow{MB})$ et les relations de Chasles du lemme 19.150(2) nous avons :

$$2(\overrightarrow{MA},\overrightarrow{MB})+(\overrightarrow{OB},\overrightarrow{OA})\in[0]_{2\pi}. \tag{19.404}$$

□

19.153.
Comment exprimer le fait qu'un angle orienté soit égal à θ modulo π alors que les angles orientés sont des classes modulo 2π ? Nous ne pouvons certainement pas écrire

$$(u,v)=[\theta]_\pi \tag{19.405}$$

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

$$[u,v]\subset[\theta]_\pi. \tag{19.406}$$

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 ente deux éléments dans $[u,v]$ est 2π alors que celle dans $[\theta]_\pi$ est π.

Si u et v forment un angle droit, nous avons

$$[u,v]=\{\frac{\pi}{2}+2k\pi\}_{k\in\mathbb{Z}}. \tag{19.407}$$

Et cela est bien un sous-ensemble de $[\pi/2]_\pi$.

Pour exprimer que la différence entre deux angles orientés diffèrent de π nous devrions écrire :

$$[u,v]\subset[a,b]_\pi \tag{19.408}$$

où le membre de droite signifie la classe modulo π d'un représentant de $[a,b]$.

Nous allons cependant nous permettre d'écrire

$$[u,v]=[a,b]_\pi \tag{19.409}$$

voire carrément

$$(u,v)=(a,b)_\pi. \tag{19.410}$$

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 alle de a à b soit ce dernier plus π.

Théorème 19.154 ([325]).
Soient 4 points distincts du plan A,B,C,D. Ils sont alignés ou cocycliques[28] *si et seulement si*

$$(\overrightarrow{CA},\overrightarrow{CB})=(\overrightarrow{DA},\overrightarrow{DB})_\pi. \tag{19.411}$$

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

$$[\overrightarrow{CA},\overrightarrow{CB}]=[\overrightarrow{DA},\overrightarrow{DB}] \tag{19.412}$$

et a fortiori l'égalité modulo π au lieu de 2π.

28. C'est-à-dire sur un même cercle.

Nous nous relâchons en termes de notations. Si les quatre points sont cocycliques, nous pouvons utiliser le théorème de l'angle inscrit 19.152 dans les triangles ABC et ADB :

$$2(\overrightarrow{CA},\overrightarrow{CB}) = (\overrightarrow{OA},\overrightarrow{OB})_{2\pi} \tag{19.413a}$$
$$2(\overrightarrow{DA},\overrightarrow{DB}) = (\overrightarrow{OA},\overrightarrow{OB})_{2\pi}, \tag{19.413b}$$

ce qui donne $2(\overrightarrow{CA},\overrightarrow{CB}) = 2(\overrightarrow{DA},\overrightarrow{DB})_{2\pi}$ et donc

$$(\overrightarrow{CA},\overrightarrow{CB}) = (\overrightarrow{DA},\overrightarrow{DB})_{\pi}. \tag{19.414}$$

Comme annoncé, nous ne faisons pas la preuve dans l'autre sens ; elle peut être trouvée dans [325]. □

Exemple 19.155
À propos de groupe engendré et de générateur[29]. Soit G le groupe des rotations d'angle[30] $k\pi/5$ (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. △

19.10.11 Angles et nombres complexes

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 24.85.

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 θ_i qu'à 2π près. Nous avons

$$[z_1, z_2] = [\theta_2 - \theta_1]_{2\pi}. \tag{19.415}$$

Soient maintenant $a, b, c, d \in \mathbb{C}$. Nous écrivons $\overrightarrow{ab}$ 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 θ_{ab} l'argument du nombre complexe $b - a$, et nous avons encore

$$[\overrightarrow{ab}, \overrightarrow{cd}] = [\theta_{ab} - \theta_{cd}]. \tag{19.416}$$

Avec toutes ces notations, ce qui est bien 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

$$\frac{a-b}{c-d} \tag{19.417}$$

est dans la classe de l'angle orienté

$$[\overrightarrow{ba}, \overrightarrow{dc}]. \tag{19.418}$$

Définition 19.156.
Soient trois points $A, O, B \in \mathbb{R}^2$. Voici comment nous définissons l'angle $\widehat{AOB}$; informelement c'est l'angle de la rotation qui fait 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 19.144).*

29. Définition 2.10 et 2.11
30. Voir la définition 19.151.

— *Soit θ l'unique élément de $[0, 2\pi[$ tel que la matrice de f dans la base canonique soit*

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{19.419}$$

par la proposition 19.138.

— *L'angle $\widehat{AOB}$ est ce nombre.*

Nous voyons que l'angle est toujours un nombre entre 0 et 2π. Par abus de notation, nous admettrons de temps en temps de parler d'angle en-dehors de cet intervalle.

Lemme-définition 19.157.
Si ℓ_1 et ℓ_2 sont deux droites de $\mathbb{R}^2$ sécantes au point O et si $x \in \ell_1$ n'est pas O, alors il existe un unique $\alpha \in [0, \pi[$ tel que $R_O(\alpha)x \in \ell_2$. La valeur de α ne dépend pas du choix du point $x \in \ell_1$.

*Cet angle α est l'**angle** de ℓ_1 à ℓ_2.*

Proposition 19.158 ([1])**.**
Les angles sont invariants sous les translations.

Plus précisément, si $A, B, S, v \in \mathbb{R}^2$, alors

$$\widehat{T_v(A)T_v(S)T_v(B)} = \widehat{ASB} \tag{19.420}$$

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 $\widehat{T_v(A)T_v(S)T_v(B)}$. □

Proposition 19.159 ([1])**.**
Les angles sont invariants par rotations, c'est à dire que si $A, B, S \in \mathbb{R}^2$ et si R_θ est une rotation, alors

$$\widehat{ASB} = \widehat{R_\theta(A)R_\theta(S)R_\theta(B)}. \tag{19.421}$$

Démonstration. Pour être plus concis, nous écrivons A_θ 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

$$\begin{cases} A'_\theta = A_\theta - S_\theta & \text{(19.422a)} \\ S'_\theta = 0 & \text{(19.422b)} \\ B'_\theta = B_\theta - S_\theta. & \text{(19.422c)} \end{cases}$$

et

$$\begin{cases} A''_\theta = \dfrac{A_\theta - S_\theta}{\|A_\theta - S_\theta\|} & \text{(19.423a)} \\ B''_\theta = \dfrac{B_\theta - S_\theta}{\|B_\theta - S_\theta\|}. & \text{(19.423b)} \end{cases}$$

Par définition, l'angle est le α tel que $R_\alpha(A''_\theta) = B''_\theta$. Nous devons prouver que le même α vérifie $R_\alpha(A'') = B''$.

Le fait que R_θ soit une isométrie nous donne déjà

$$\|R_\theta(A) - R_\theta(B)\| = \|A - B\|. \tag{19.424}$$

Ensuite, la relation de définition de α s'écrit

$$\frac{R_\alpha R_\theta(A) - R_\alpha R_\theta(S)}{\|A_\theta - S_\theta\|} = \frac{R_\theta(B) - R_\theta(S)}{\|B_\theta - S_\theta\|}. \tag{19.425}$$

Vu que R_α et R_θ commutent, nous avons

$$R_\theta \frac{R_\alpha(A) - R_\alpha(B)}{\|A - S\|} = R_\theta \frac{B - S}{\|B - S\|}, \tag{19.426}$$

et comme R_θ est inversible, cela donne $R_\alpha(A'') = B''$. □

Lemme 19.160 ([1]).
Soit $A \in \mathbb{R}^2$ et une droite ℓ_1. Soit ℓ_2 une droite passant par A et intersectant ℓ_1 en O. Alors

$$\sigma_{\ell_1}(A) = R_O(-2\alpha)A \tag{19.427}$$

où α est l'angle de ℓ_1 à ℓ_2.

Démonstration. Nous allons utiliser des coordonnées autour de O. Il existe un vecteur v tel que

$$A = O + v \tag{19.428}$$

Par définition de l'angle α[31], la droite ℓ_2 s'obtient par rotation d'angle α depuis la droite ℓ_1. Donc le point

$$B = R_O(-\alpha)A \tag{19.429}$$

est sur ℓ_1.

Nous allons prouver que le point

$$D = R_O(-2\alpha)A \tag{19.430}$$

est $D = \sigma_{\ell_1} A$.

Nous commençons par montrer que la droite (DA) est perpendiculaire à ℓ_1, c'est-à-dire que

$$(D - A) \cdot (B - O) = 0. \tag{19.431}$$

En utilisant le fait que

$$R_O(\alpha)(O + X) = O + R(\alpha)X, \tag{19.432}$$

nous avons

$$D - A = R_O(-2\alpha)(O + v) - (O + v) = O + R(-2\alpha)v - O - v = R(-2\alpha)v - v \tag{19.433}$$

et de la même façon,

$$B - O = R(-\alpha)v. \tag{19.434}$$

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

$$(D - A) \cdot (B - O) = \langle R(-2\alpha)v - v, R(-\alpha)v \rangle \tag{19.435a}$$
$$= \langle R(-\alpha)v - R(\alpha)v, v \rangle. \tag{19.435b}$$

En utilisant la matrice de rotation du lemme 19.135 nous trouvons

$$\big(R(-\alpha) - R(\alpha)\big)v = \begin{pmatrix} 2\sin(\alpha)v_2 \\ -2\sin(\alpha)v_1 \end{pmatrix} \tag{19.436}$$

et donc

$$\big\langle \big(R(-\alpha) - R(\alpha)\big)v, v \big\rangle = 0. \tag{19.437}$$

Le point D est bien sûr la droite perpendiculaire à ℓ_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 $\sigma_{\ell_1}(A)$. Vu que D n'est pas A, nous avons $D = \sigma_\ell(A)$. □

31. Définition 19.157.

19.10.12 Classification

Théorème 19.161 ([313]).
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.

Si f n'a pas de points fixes Soit $x \in \mathbb{R}^2$. Nous considérons le segment $[f, f(x)]$ et nous nommons l sa médiatrice. Par construction, $f(x) = \sigma_l(x)$. Nous posons $g = \sigma_l \circ f$, et nous avons

$$g(x) = x. \tag{19.438}$$

Donc nous avons $f = \sigma_l \circ g$ avec $x \in \mathrm{Fix}(g)$.

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

$$g(y) = y \tag{19.439}$$

et $g(x) = x$ parce que

$$d\big(x, f(y)\big) = d\big(f(x), f(y)\big) = d(x, y), \tag{19.440}$$

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 et donc toute la droite (xy).

Si f fixe une droite 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

$$\begin{cases} d\big(x, f(z)\big) = d(x, z) & \text{(19.441a)} \\ d\big(y, f(z)\big) = d(y, z) & \text{(19.441b)} \end{cases}$$

ce qui signifie que $f(z)$ est sur l'intersection des deux cercles[32] $S\big(x, d(x, z)\big)$ et $S\big(y, d(y, z)\big)$, et comme ce sont deux cercles centrés sur la droite l, les intersections sont liées par σ_l. Autrement dit, les intersections sont z et $\sigma_l(z)$.
Si $f(z) = z$ alors f fixe trois points non alignés et fixe dont $\mathbb{R}^2$, c'est-à-dire $f = \mathrm{Id}$.
Si par contre $f(z) = \sigma_l(z)$ alors les isométries f et σ_l coïncident sur trois points et coïncident donc partout par le corollaire 19.6 : $f = \sigma_l$.

Conclusion Nous avons montré que si $\mathrm{Fix}(f)$ a dimension m, alors il existe une droite pour laquelle $f = \sigma_l \circ g$ avec $\dim\big(\mathrm{Fix}(g)\big) > m$. Donc il faux au maximum trois pas pour avoir $\dim\big(\mathrm{Fix}(g)\big) = 2$ c'est-à-dire pour avoir $g = \mathrm{Id}$.

□

Définition 19.162.
*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 ℓ.*

Théorème 19.163 ([313]).
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 translations d'axes parallèles),
(4) les rotations (composées de 2 réflexions d'axes non parallèles),
(5) les réflexions glissées (composées de 3 réflexions)

32. L'intersection existe pare que $d(x, z) + d(y, z) > d(x, y)$.

Démonstration. Nous savons déjà que $f \in \mathrm{Isom}(\mathbb{R}^2)$ est une composée de 0, 1, 2 ou 3 réflexions.

Zéro réflexions Alors c'est l'identité. Ce n'est pas très profond.

Une réflexion Alors f est une réflexion. Toujours pas très profond.

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 ℓ_1 tel que $\ell_1 + v = \ell_2$. Nous allons utiliser le lemme 19.114 pour montrer que $\sigma_{\ell_1} \circ \sigma_{\ell_2} = \tau_{2v}$. Nous avons

$$\ell_1 = \ell_0 + w \tag{19.442a}$$
$$\ell_2 = \ell_0 + w + v \tag{19.442b}$$

où w est un vecteur perpendiculaire à ℓ_1 et ℓ_0 est la droite passant par l'origine et parallèle à ℓ_1 et ℓ_2. Avec cela,

$$\begin{aligned}(\sigma_{\ell_1} \circ \sigma_{\ell_2})(x) &= \sigma_{\ell_1}\big(\sigma_{\ell_0}(x) + 2w\big) && (19.443a)\\ &= \sigma_{\ell_0}\big(\sigma_{\ell_0}(x) + 2w\big) + 2(v + w) && (19.443b)\\ &= x + \underbrace{\sigma_{\ell_0}(2w)}_{-2w} + 2v + 2w && (19.443c)\\ &= x + 2v. && (19.443d)\end{aligned}$$

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 ℓ_1 et ℓ_2 non parallèles. Nous notons O le point d'intersection, et nous allons voir que $f = R_O(2\alpha)$ où α est l'angle de ℓ_1 à ℓ_2 donné par le lemme 19.157.

Soit $x \in \ell_1$. Alors

$$f(x) = \sigma_{\ell_2}(x), \tag{19.444}$$

et le lemme 19.160 nous donne un moyen de calculer $\sigma_{\ell_2}(x)$ parce que ℓ_1 est une droite passant par x et coupant ℓ_1 au point O. Le lemme dit que $\sigma_{\ell_2}(x) = R_O(2\alpha)$. Remarque : c'est bien 2α et non -2α parce qu'il s'agit de l'angle de ℓ_2 à ℓ_2 ; 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

$$f(y) = \sigma_{\ell_2}\big(R_O(-2\alpha)y\big) \tag{19.445}$$

Nous posons $z = \sigma_{\ell_1}(y) = R_O(-2\alpha)y$. Soit la droite ℓ_3 passant par O et z. Vu que $R_O(2\alpha)z = y \in \ell_2$, l'angle de ℓ_3 à ℓ_2 est 2α. Par conséquent

$$\sigma_{\ell_2}(z) = R_O\big(-2 \times (-2\alpha)\big)z = R_O(4\alpha)z = R_O(4\alpha)R_O(-2\alpha)y = R_O(2\alpha)y. \tag{19.446}$$

Donc les transformations f et $R_O(2\alpha)$ coïncident pour tous les points des droites ℓ_1 et ℓ_2, qui ne sont pas parallèles. Cela prouve que $f = R_O(2\alpha)$.

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 :

(1) $f = \sigma_\ell \circ \tau_v$,
(2) $f = \tau_v \circ \sigma_\ell$,
(3) $f = \tau_v \circ \sigma_\ell$ avec $v \parallel \ell$.

À chacune de ces étapes, v et ℓ 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 (2).

Nous supposons que ℓ_2 n'est pas parallèle à ℓ_3. Dans ce cas, si $O = \ell_2 \cap \ell_3$ nous avons

$$\sigma_{\ell_3} \circ \sigma_{\ell_2} = R_O(2\alpha) \tag{19.447}$$

où α est l'angle de ℓ_2 à ℓ_3. En réalité tant que l'angle de ℓ_3' à ℓ_2' est α nous avons

$$\sigma_{\ell_3'} \circ \sigma_{\ell_2'} = \sigma_{\ell_3} \circ \sigma_{\ell_2} = R_O(2\alpha). \tag{19.448}$$

Nous choisissons ℓ_2' parallèle à ℓ_1, de telle sorte à ce que $\sigma_{\ell_2'} \circ \sigma_{\ell_1}$ soit une translation. Alors nous avons

$$f = \sigma_{\ell_3} \circ \sigma_{\ell_2} \circ \sigma_{\ell_1} = \sigma_{\ell_3} \circ \sigma_{\ell_2'} \circ \sigma_{\ell_1'} = \sigma_{\ell_3} \circ \tau_v. \tag{19.449}$$

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 (1) soit sous la forme (2).

Nous prouvons à présent que toute transformation de la forme (1) peut être écrite sous la forme (2). Plus précisément nous allons prouver que si ℓ est une droite, v un vecteur et ℓ_0 la droite parallèle à ℓ passant par l'origine, alors

$$\sigma_\ell \circ \tau_v = \tau_{\sigma_{\ell_0}(v)} \circ \sigma_l \tag{19.450}$$

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 σ_{ℓ_0} est linéaire :

$$(\sigma_\ell \circ \tau_v)(x) = \sigma_l(x + v) = \sigma_{\ell_0}(x) + \sigma_{\ell_0}(v) + 2w, \tag{19.451}$$

et

$$(\tau_{\sigma_{\ell_0}(v)} \circ \sigma_\ell)(x) = \sigma_{\ell_0}(v) + \sigma_\ell(x) = \sigma_{\ell_0}(v) + \sigma_{\ell_0}(x) + 2w. \tag{19.452}$$

L'égalité est faite.

Nous montrons maintenant que toute transformation de la forme (2) peut être mise sous la forme (3). Soit donc $f = \tau_v \circ \sigma_\ell$ où v et ℓ 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 ℓ_0 la droite parallèle à ℓ et ℓ' et passant par l'origine. Soit aussi le vecteur w tel que $\ell = \ell_0 + w$. Alors nous avons

$$\begin{cases} \sigma_\ell = \sigma_{\ell_0} + 2w & \text{(19.453a)} \\ \sigma_{\ell'} = \sigma_{\ell_0} + 2w + v_1 & \text{(19.453b)} \end{cases}$$

Nous avons

$$(\tau_v \circ \sigma_\ell)(x) = v + \sigma_{\ell_0}(x) + 2w \tag{19.454}$$

et

$$\begin{aligned} (\tau_{v_2} \circ \sigma_{\ell'})(x) &= v_2 + \sigma_{\ell_0}(x) + 2w + v_1 && \text{(19.455a)} \\ &= \sigma_{\ell_0}(x) + v + 2w && \text{(19.455b)} \end{aligned}$$

où dans la dernière ligne, nous avons regroupé $v_1 + v_2 = v$. Et voilà.

□

19.10.13 Sous-groupe fini d'isométries du plan

Théorème 19.164 ([326]).
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 réflexions, alors il est cyclique*[33] *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 σ est une réflexion dans G, alors nous avons $G = \text{gr}\big(\sigma, R_C(4\pi/n)\big)$ où $R_C(\theta)$ est la rotation d'angle θ 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.

Pour (1) C'est la proposition 19.86.

Questions de réflexions Le théorème 19.161 nous dit que les éléments de G sont des compositions d'au maximum 3 réflexions.

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 19.163 sont des réflexions glissées[34], c'est à dire des transformations de la forme $g = \tau_v \circ \sigma_\ell$ où v est un vecteur parallèle à la droite ℓ. Si $x \in \ell$, alors

$$g(x) = \tau_v(x) = x + v, \tag{19.456}$$

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.

G^+ et G^- Pour la même raison que celle qui exclu les réflexions glissées, G ne peut pas contenir de translations. Le théorème 19.163 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.

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

$$\begin{aligned} \varphi\colon G^+ &\to G^- \\ f &\mapsto \sigma \circ f \end{aligned} \tag{19.457}$$

est une bijection.

Surjective Soit $s \in G^-$. Posons $f = \sigma^{-1} \circ s$. Vu que σ^{-1} et s sont des réflexions, f est une rotation. Donc $f \in G^+$ et $\varphi(f) = s$.

Injective La condition $\varphi(f) = \varphi(g)$ dit que $\sigma \circ f = \sigma \circ g$. En composant par σ^{-1} nous obtenons $f = g$.

$G = \text{gr}\big(R_C(2\pi/p)\big)$ 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 corollaire 2.32 au théorème de 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

33. Définition 2.11.
34. Définition 19.162.

$g^p = \text{Id}$. Or il n'y a que p telles rotations. Donc elles sont toutes dans G^+. Nous en déduisons que

$$G^+ = \text{gr}\big(R_C(2\pi/p)\big). \tag{19.458}$$

Pour (2) Dans le cas où G ne contient pas de réflexions, G^- est vide et G contient n éléments. La relation (19.458) devient

$$G = G^+ = \text{gr}\big(R_C(2\pi/n)\big). \tag{19.459}$$

Pour (3) Nous supposons maintenant que G contienne au moins une réflexion. De la sorte $G^- \neq \varnothing$.

Pour (3a) Les seuls points fixes d'une réflexions sont ceux de l'axe. Donc C soit 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.

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 19.457, si G^- est non vide, G possède un nombre pair d'éléments.

Pour (3c) Si $\sigma \in G$ est une réflexion, nous savons que G^+ possède $p = n/2$ éléments et que

$$G^+ = \{R_C(2k\pi/p)\} = \{R_C(4k\pi/n)\}_{k=1,\dots,n/2}. \tag{19.460}$$

L'élément $\sigma \in G^-$ étant fixé, la bijection (19.457) nous indique que tous les éléments de G^- sont de la forme $\sigma \circ f$ avec $f \in G^+$. Donc

$$G^- \subset \text{gr}\big(\sigma, R_C(4\pi/n)\big). \tag{19.461}$$

Nous avons aussi

$$G^+ \subset \text{gr}\big(\sigma, R_C(4\pi/n)\big). \tag{19.462}$$

Et comme σ et $R_C(4\pi/n)$ sont dans G nous avons $\text{gr}\big(\sigma, R_C(4\pi/n)\big) \subset G$. Tout cela pour dire que

$$G = \text{gr}\big(\sigma, R_C(4\pi/n)\big). \tag{19.463}$$

$R\sigma = \sigma R^{-1}$ Nous restons dans le cas où G^- n'est pas vide. Nous considérons R, la rotation d'angle θ autour de C. Si R_0 est la rotation d'angle θ autour de $(0,0)$, nous avons

$$R = \tau_C \circ R_0 \circ \tau_C^{-1}, \tag{19.464}$$

et si σ_0 est la symétrie d'axe parallèle à l'axe de σ, mais passant par $(0,0)$ nous avons :

$$\sigma = \tau_C \circ \sigma_0 \circ \tau_C^{-1}. \tag{19.465}$$

Si v est le vecteur directeur de la réflexion σ_0, nous considérons enfin α, la rotation qui fait $\alpha(v) = (1,0)$. Nous avons alors

$$\sigma_0 = \alpha^{-1} \circ s \circ \alpha \tag{19.466}$$

où s est la symétrie autour de l'axe horizontal. En n'ayant pas peur d'identifier $\mathbb{R}^2$ à $\mathbb{C}$, l'applicaiton s est la conjugaison complexe. Avec tout ça nous avons

$$R\sigma = \tau_C R_0 \tau_C^{-1} \tau_C \sigma_0 \tau_C^{-1} = \tau_C R_0 \alpha^{-1} s \alpha \tau_C^{-1} = \tau_C \alpha^{-1} R_0 s \alpha \tau_C^{-1} \tag{19.467}$$

où nous avons utilisé le fait que les rotations autour de $(0,0)$ forment un groupe abélien pour commuter α^{-1} avec R_0. Nous utilisaons à présent le lemme 19.97 pour commuter R avec s :

$$\begin{aligned}
R\sigma &= \tau_X \alpha^{-1} s R_0^{-1} \alpha \tau_C^{-1} && \text{(19.468a)}\\
&= \tau_C \underbrace{\alpha^{-1} s \alpha}_{\sigma_0} R_0^{-1} \tau_C^{-1} && \text{(19.468b)}\\
&= \tau_C \sigma_0 \tau_C^{-1} \tau_C R_0^{-1} \tau_C^{-1} && \text{(19.468c)}\\
&= \sigma R^{-1}. && \text{(19.468d)}
\end{aligned}$$

Nous avons utilisé le fat que $\tau_C R_0^{-1} \tau_C^{-1} = R^{-1}$ comme on peut s'en convaincre en calculant le produit.

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 σR où R est une rotation autour de C. Pour savoir la table de multiplication de G, nous devons écrire

$$(\sigma^{\epsilon_1} R^k)(\sigma^{\epsilon_2} R^l) = \sigma^{\epsilon} R^m \tag{19.469}$$

où $\epsilon_1, \epsilon_2 \in \{0, 1\}$, R est la rotation d'angle $4\pi/n$ autour de C et α et m sont des constantes à exprimer en fonction de ϵ_1, ϵ_2, k et l.

Tous les éléments de G pouvant être écrits soit sous la forme R^m soit sous la forme σR^m, nous avons les possibilités suivantes :

(1) $R^m R^l = R^{m+l}$

(2) $(R^m)(\sigma R^l) = \sigma R^{-m} R^l = \sigma R^{l-m}$

(3) $(\sigma R^m) R^l = \sigma R^{m+l}$

(4) $(\sigma R^m)(\sigma R^l) = \sigma\sigma R^{l-m} = R^{l-m}$.

Pour (3d) Récoltons quelque faits.

— 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 σR^m suivant que l'élément soit 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 19.94(2)).

L'application $\varphi : G \to D_n$ suivante est donc une bijection :

$$\begin{cases} \varphi(R^m) = r^m & \text{(19.470a)} \\ \varphi(\sigma R^m) = sr^m. & \text{(19.470b)} \end{cases}$$

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 19.98 et celle du groupe G que nous venons de faire.

□

Définition 19.165 (Groupe de symétrie d'une partie de $\mathbb{R}^n$[313])**.**
Si Y est une partie de $\mathbb{R}^n$, nous définissons le ***groupe des symétries*** *de Y par*

$$\mathrm{Sym}(Y) = \{f \in \mathrm{Isom}(\mathbb{R}^n) \text{ tel que } f(Y) = Y\}. \tag{19.471}$$

Nous définissons aussi le groupe des symétries propres de Y par

$$\mathrm{Sym}^+(Y) = \{f \in \mathrm{Isom}^+(\mathbb{R}^n) \text{ tel que } f(Y) = Y\}. \tag{19.472}$$

Théorème 19.166 ([313])**.**
Soit $Y \subset \mathbb{R}^2$ tel que le groupe $\mathrm{Sym}^+(Y)$ soit fini d'ordre n. Alors c'est un groupe cyclique d'ordre n.

Si $\mathrm{Sym}^+(Y)$ est fini, alors $\mathrm{Sym}(Y)$ est soit cyclique[35] *d'ordre n soit isomorphe au groupe diédral*[36] *d'ordre $2n$.*

Démonstration. Nous savons déjà par la proposition 19.87 que $\mathrm{Sym}^+(Y)$ est isomorphe à un sous-groupe H^+ d'ordre n de $\mathrm{SO}(2)$. Vérifions que ce groupe est cyclique. Si $n = 1$, c'est évident. Si $n \geqslant 2$ 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 α_0 cet angle.

35. Définition 2.11.
36. Définition 19.89.

Nous montrons que H^+ est engendré par la rotation d'angle α_0. Soit une rotation d'angle α. Étant donné que $\alpha_0 < \alpha$ nous pouvons effectuer la division euclidienne[37] de α par α_0 et obtenir

$$\alpha = k\alpha_0 + \beta \tag{19.473}$$

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 α est un multiple de α_0, ce qui signifie que le groupe H^+ est cyclique engendré par α_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 19.87 nous avons un homomorphisme injectif

$$\phi\colon \text{Sym}(Y) \to \text{O}(2), \tag{19.474}$$

et en posant $H = \phi\big(\text{Sym}(Y)\big)$ nous avons un isomorphisme de groupes $\phi\colon \text{Sym}(Y) \to H$. Nous savons aussi que ce ϕ se restreint en

$$\phi\colon \text{Sym}^+(Y) \to H^+ \subset \text{SO}(2) \tag{19.475}$$

où $H^+ = \phi\big(\text{Sym}^+(Y)\big) = H \cap \text{SO}(2)$. 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 ϕ 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 19.163 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 réflexions glissées 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 σ. 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

$$h = (h\sigma^\epsilon)\sigma^\epsilon \tag{19.476}$$

pour n'importe quelle valeur de ϵ, et en particulier pour $\epsilon = \pm 1$.

Si $h \in \text{SO}(2)$ alors nous écrivons $h = h\epsilon^0$ 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

$$H = H^+C_2. \tag{19.477}$$

Cette décomposition est unique parce que si $h_1c_1 = h_2c_2$ alors $h_2^{-1}h_1 = c_2c_1^{-1}$, et comme $h_2^{-1}h_1 \in H^+$ nous avons $c_2c_1^{-1} \in H^+$ et donc $c_1 = c_2$. Partant nous avons aussi $h_1 = h_2$. Pour avoir le produit semi-direct il faut encore montrer que $\mathbf{Ad}(C_2)H^+ \subset H^+$. Le seul cas à vérifier est $\mathbf{Ad}(\sigma)H^+ \subset H^+$. Vu que les éléments de H^+ sont caractérisés par le fait d'avoir un déterminant positif, nous avons

$$\mathbf{Ad}(\sigma)R(\alpha) = \sigma R(\alpha)\sigma^{-1} \in H^+. \tag{19.478}$$

□

Remarque 19.167.
Tout ceci est cohérent avec le théorème de Burnside 11.262 parce que le sous-groupe fini de $\text{SO}(n)$ engendré par la rotation $R(2\pi/n)$ est un groupe d'exposant fini, à savoir que si h est dans ce groupe, $h^n = \text{Id}$.

37. Théorème 3.6.

19.10.14 Relations trigonométriques dans un triangle rectangle

Nous donnons maintenant quelque relations trigonométriques classiques dans un triangle rectangle. Le théorème de Pythagore est déjà le théorème 11.27 ; nous nous concentrons ici sur les angles.

Proposition 19.168.
Soient $A, B, S \in \mathbb{R}^2$ des points distincts et non alignés formant un triangle rectangle en A :

$$(A - S) \cdot (B - A) = 0. \tag{19.479}$$

En posant $\theta = \widehat{ASB}$ nous avons

$$\cos(\theta) = \frac{\|A - S\|}{\|B - S\|} \tag{19.480}$$

et

$$\sin(\theta) = \pm\frac{\|B - A\|}{\|B - S\|}. \tag{19.481}$$

Démonstration. Nous posons $C = A - S$ et $D = B - S$. Vu que $C \neq 0$, il existe une rotation R_α telle que

$$\begin{cases} (R_\alpha C)_x > 0 & \text{(19.482a)} \\ (R_\alpha C)_y = 0. & \text{(19.482b)} \end{cases}$$

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

$$X \cdot (Y - X) = 0. \tag{19.483}$$

De ce fait, le point Y satisfait à $Y_x = X_x$. Et enfin, grâce aux propositions 19.158 et 19.159 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

$$\begin{cases} \cos(\theta) = X'_x Y'_x & \text{(19.484a)} \\ \sin(\theta) = X'_x Y'_y. & \text{(19.484b)} \end{cases}$$

Vu que $X = (X_x, 0)$ nous avons $X'_x = 1$. De plus

$$\|Y\| = \|R_\alpha(D)\| = \|D\| = \|B - S\|. \tag{19.485}$$

En substituant les valeurs dans (19.484),

$$\cos(\theta) = Y'_x = \frac{Y_x}{\|Y\|} = \frac{X_x}{\|B - S\|} = \frac{\|C\|}{\|B - S\|} = \frac{\|A - S\|}{\|B - S\|}. \tag{19.486}$$

Voila 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

$$\sin(\theta) = Y'_y = \frac{Y_y}{\|Y\|} = \frac{\pm\|Y - X\|}{\|Y\|} = \pm\frac{\|D - C\|}{\|Y\|} = \pm\frac{\|B - A\|}{\|B - S\|}. \tag{19.487}$$

□

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 19.169.
Soient trois points non alignés $A, S, B \in \mathbb{R}^2$. Nous avons

$$\cos\left(\widehat{ASB}\right) = \frac{(A-S)\cdot(B-S)}{\|A-S\|\|B-S\|}. \tag{19.488}$$

Démonstration. Nous posons $C = A - S$, $D = B - S$, $X = C/\|C\|$ et $Y = D/\|D\|$. Avec cela, la définition 19.156 donne l'équation

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} X_x \\ X_u \end{pmatrix} = \begin{pmatrix} Y_x \\ Y_y \end{pmatrix} \tag{19.489}$$

que nous écrivons comme le système

$$\begin{cases} X_c\cos(\theta) - X_y\sin(\theta) = Y_x & \text{(19.490a)} \\ X_y\cos(\theta) + X_x\sin(\theta) = Y_y. & \text{(19.490b)} \end{cases}$$

Nous considérons maintenant cela comme un système pour $\big(\cos(\theta), \sin(\theta)\big)$:

$$\begin{pmatrix} X_x & -X_y \\ X_y & X_x \end{pmatrix} \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} Y_x \\ Y_y \end{pmatrix}. \tag{19.491}$$

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

$$\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \begin{pmatrix} X_x & X_y \\ -X_y & X_x \end{pmatrix} \begin{pmatrix} Y_x \\ Y_y \end{pmatrix}. \tag{19.492}$$

Cela donne ce que nous voulions :

$$\cos(\theta) = X\cdot Y = \frac{C\cdot D}{\|C\|\|D\|} = \frac{(A-S)\cdot(B-S)}{\|A-S\|\|C-S\|}. \tag{19.493}$$

□

Remarque 19.170.
Prendre la formule (19.488) 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 prennent ça pour définition d'angle 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 $\mathrm{SO}(2)$ et aux fonctions trigonométriques.

19.10.15 Pavages du plan

Définition 19.171.
*Une application affine $f : \mathbb{R}^n \to \mathbb{R}^n$ est un **déplacement** lorsqu'elle est une isométrie de $(\mathbb{R}^n, d)$ qui préserve l'orientation*[38].

Définition 19.172 ([234]).
*Un **pavage** de $\mathbb{R}^2$ est une paire (G, K) où G est un groupe de déplacements*[39] *de $\mathbb{R}^2$ et K un compact de $\mathbb{R}^2$ d'intérieur non vide telle que*

(1) $G\cdot K = \mathbb{R}^2$,

(2) Si $g_1, g_2 \in G$ satisfont $g_1\cdot\mathrm{Int}(K)\cap g_2\cdot\mathrm{Int}(K) \neq 0$, alors $g_1\cdot K = g_2\cdot K$.

*Nous disons qu'un groupe 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.*

38. Définition 11.70.
39. Définition 19.171.

En termes de notations,

$$G \cdot K = \bigcup_{g\in G} g(K) = \bigcup_{g\in G}\bigcup_{k\in K} g(k). \tag{19.494}$$

Lemme 19.173 ([1]).
Soient une bijection affine $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^2$ 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 $f\colon \mathbb{R}^2 \to \mathbb{R}$ et la définition

$$f = \{x \in \mathbb{R}^2 \text{ tel que } f(x) = 0\}. \tag{19.495}$$

Nous supposons que $f(A) > 0$; sinon, nous pouvons utiliser $-f$ au lieu de f. Donc

$$S = \{x \in \mathbb{R}^2 \text{ tel que } f(x) > 0\}. \tag{19.496}$$

La partie $\varphi(S)$ est alors donnée par

$$\varphi(S) = \{\varphi(x) \text{ tel que } f(x) > 0\}. \tag{19.497}$$

Comme φ est une bijection, cela s'écrit aussi bien

$$\varphi(S) = \{y \in \mathbb{R}^2 \text{ tel que } f\big(\varphi^{-1}(y)\big) > 0\}. \tag{19.498}$$

De même

$$\varphi(d) = \{s \in \mathbb{R}^2 \text{ tel que } f\big(\varphi^{-1}(s)\big) = 0\}. \tag{19.499}$$

Donc les deux côtés de la droite $\varphi(d)$ sont donnés par le signe de $f \circ \varphi^{-1}$. Nous avons

$$(f \circ \varphi^{-1})\big(\varphi(A)\big) = f(A) > 0. \tag{19.500}$$

Donc $\varphi(A) \in \varphi(S)$. □

Lemme 19.174.
Le groupe

$$G = \mathrm{gr}(\tau_{e_1}, \tau_{e_2}) \tag{19.501}$$

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 intersections des intérieurs des carrés. Notons toutefois qu'il y a un recouvrement des bords. □

Lemme 19.175 ([1, 327]).
Le groupe

$$G = \mathrm{gr}(\tau_{e_1}, \tau_{e_2}, R_0(\pi)) \tag{19.502}$$

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 19.9.

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 19.172 d'un pavage.

Supposons que $g_1 \cdot \mathrm{Int}(K) \cap g_2 \cdot \mathrm{Int}(K) \neq \varnothing$. Cela signifie qu'il existe $k_1, k_2 \in \mathrm{Int}(K)$ tels que $g_1(k_1) = g_2(k_2)$, ou encore que

$$(g_2^{-1}g_1)k_1 = k_2 \in \mathrm{Int}(K). \tag{19.503}$$

Quel est la forme d'un élément général de G ? Le lemme 2.8 nous indique qu'un élément général de G est un produit fini de τ_{e_1}, τ_{e_2} et $R_0(\pi)$. Mais nous savons que si α est linéaire,

$$\alpha \circ \tau_u = \tau_{\alpha(u)} \circ \alpha. \tag{19.504}$$

Dans notre cas, dans un produit général, nous pouvons déplacer tous les facteurs $R_0(\pi)$ à droite en changeant des τ_{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

$$g = \tau_{e_1}^k \tau_{e_2}^l R_0(\pi)^m \tag{19.505}$$

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 (19.503) doivent donc vérifier $g_2^{-1}g_1 = \text{Id}$, et donc $g_1 = g_2$. Par conséquent $g_1 \cdot K = g_2 \cdot K$. □

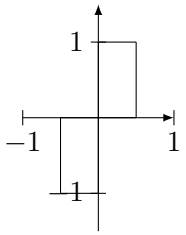


FIGURE 19.9 – Le compact K et son image par $R_0(\pi)$ pour le lemme 19.175.

Lemme 19.176 ([1, 327])**.**
Le groupe

$$\text{gr}(\tau_{e_1}, \tau_{e_2}, R_0(\pi/2)) \tag{19.506}$$

est un groupe de pavage.

Démonstration. Le pavé à considérer est

$$K = [-\frac{1}{2}, 0] \times [0, \frac{1}{2}]. \tag{19.507}$$

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 19.175. 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 τ_{e_1}, τ_{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

$$g = \tau_a \circ R_0(\pi/2)^k \tag{19.508}$$

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 nombre 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 τ_a ne permet d'avoir $(\tau_a \circ R_O(\pi/2))k \in \text{Int}(K)$.

Bref, si $(g_2^{-1}g_2)k \in K$, alors encore une fois $g_2^{-1}g_1 = \text{Id}$. □

Et c'est maintenant que les choses compliquées commencent.

Lemme 19.177.
Le groupe

$$\text{gr}\left(\tau_{e_1}, \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(\pi/3)\right) \tag{19.509}$$

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 $f_{AB}(x,y) = x - \sqrt{3}y$. Vu que $f_{AB}(C) = 1$, la première inéquation pour K est

$$x - \sqrt{3}y \geqslant 0. \tag{19.510}$$

— 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

$$-x - \sqrt{3}y + 1 \geqslant 0 \tag{19.511}$$

— 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 :

$$y \geqslant 0. \tag{19.512}$$

En résumé, la définition de K est le système

$$\begin{cases} x - \sqrt{3}y \geqslant 0 & \text{(19.513a)} \\ -x - \sqrt{3}y + 1 \geqslant 0 & \text{(19.513b)} \\ y \geqslant 0. & \text{(19.513c)} \end{cases}$$

La figure 19.10(a) nous montre ce triangle et l'action des puissances de $R_0(\pi/3)$ sur lui. Pour votre lanterne, la matrice de cette rotation est

$$R_0(\pi/3) = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}. \tag{19.514}$$

La figure 19.10(c) vous montre une partie de ce pavage dans toute sa splendeur.

Vu que 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 19.10(b).

Pour montrer que les triangles dessinés pavent effectivement le parallélogramme, nous allons faire 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 $x + \frac{1}{\sqrt{3}}y - 1 \leqslant 0$ et les autres.

Si $x + y/\sqrt{3} - 1 \leqslant 0$ 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 : $x \geqslant \frac{1}{2}$ et $x \leqslant \frac{1}{2}$.

Si $x \leqslant \frac{1}{2}$ Enfin nous coupons avec la droite diagonale partant de $(0,0)$, c'est à dire selon que $x - \sqrt{3}y \leqslant 0$ ou $x - \sqrt{3}y \geqslant 0$.

Si $x - \sqrt{3}y \leqslant 0$ Les points dont nous parlons sont les $(x,y) \in \mathbb{R}^2$ vérifiant

$$\begin{cases} x + \dfrac{1}{\sqrt{3}}y - 1 \leqslant 0 & \text{(19.515a)} \\ x \leqslant \dfrac{1}{2} & \text{(19.515b)} \\ x - \sqrt{3}y \leqslant 0. & \text{(19.515c)} \end{cases}$$

En suivant le dessin, vous remarquerez que l'élément de G à considérer est

$$g = \tau_{e_1} \circ \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})} \circ R_0(\pi/3)^4. \tag{19.516}$$

Nous avons

$$g(A) = (\frac{1}{2}, \frac{\sqrt{3}}{2}) \tag{19.517a}$$
$$g(B) = (\frac{1}{2}, \frac{\sqrt{3}}{6}) \tag{19.517b}$$
$$g(C) = (0, 0). \tag{19.517c}$$

Ce que les équations (19.515) décrivent est l'intersection des trois parties suivanteas[40] :

— le côté de la droite $\big(g(A)g(B)\big)$ contenant $g(C)$,

— le côté de la droite $\big(g(A)g(C)\big)$ contenant $g(B)$,

— le côté de la droite $\big(g(B)g(C)\big)$ contenant $g(A)$,

Le lemme 19.173 nous permet de d'exprimer ces trois parties en termes de K :

— l'image par g du côté de la droite (AB) content C,

— l'image par g du côté de la droite (AC) content B,

— l'image par g du côté de la droite (BC) content A,

Comme g est une bijection, l'intersection des images par g et l'image par g de l'intersection. Bref, les équations (19.517) décrivent l'image par g de K.

Cette partie est donc pavée par (G, K).

Si $x - \sqrt{3}y \geqslant 0$ hop.

Si $x \geqslant \frac{1}{2}$ hop.

Si $x + y/\sqrt{3} - 1 \geqslant 0$ hop.

Les cas listés se traitement surement 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 19.172. 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 lemme 2.8 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 τ_{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 :

$$R_0(\pi/3)^k e_1 \tag{19.518a}$$
$$R_0(\pi/3)^k (\frac{1}{2}, \frac{\sqrt{3}}{2}) \tag{19.518b}$$
$$R_0(\pi/3)^{-k} e_1 \tag{19.518c}$$
$$R_0(\pi/3)^{-k} (\frac{1}{2}, \frac{\sqrt{3}}{2}). \tag{19.518d}$$

En utilisant la matrice (19.514) nous trouvons assez vite que

$$R_0(\pi/3)e_1 = \begin{pmatrix} 1/2 \\ \sqrt{3}/2 \end{pmatrix} \tag{19.519a}$$
$$R_0(\pi/3)^2 e_1 = \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix} \tag{19.519b}$$
$$R_0(\pi/3)^3 e_1 = -e_1. \tag{19.519c}$$

40. Personnellement, je n'ai pas vérifié, mais ça m'étonnerait que ce soit faux. Vérifiez et écrivez-moi.

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 $\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right)$ sont déjà parmi celles listées en (19.519). 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

$$\tau_v\circ R_0(\pi/3)^k \tag{19.520}$$

avec v dans

$$\mathbb{Z}e_1+\mathbb{Z}\begin{pmatrix}1/2\\ \frac{\sqrt{3}}{2}\end{pmatrix}+\mathbb{Z}\begin{pmatrix}-1/2\\ \frac{\sqrt{3}}{2}\end{pmatrix}. \tag{19.521}$$

Remarquons que

$$\begin{pmatrix}-1/2\\ \sqrt{3}/2\end{pmatrix}=-e_1+\begin{pmatrix}1/2\\ \sqrt{3}/2\end{pmatrix}. \tag{19.522}$$

Donc un élément général de G est

$$\tau_{e_1}^k\circ\tau^l_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}\circ R_0(\pi/3)^m. \tag{19.523}$$

Nous pouvons maintenant prouver notre point. Pour cela nous allons successivement considérer les 5 rotations de K présentées dans la sous-figure 19.10(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.

Pour $L=R_0(\pi/3)K$

Pour $L=R_0(\pi/3)^2K$ Nous allons prouver que si $(x,y)\in\text{Int}(L)$, alors

$$\tau_{e_1}^k\circ\tau^l_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}(x,y) \tag{19.524}$$

ne peut pas être dans $\text{Int}(K)$. Pour la simplicité des notations nous notons $r=R_0(\pi/3)^2$; nous avons

$$r=\begin{pmatrix}-1/2 & -\sqrt{3}/2\\ \sqrt{3}/2 & -1/2\end{pmatrix}. \tag{19.525}$$

Nous considérons $g=\tau_{e_1}^k\circ\tau_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}\circ r$. Un calcul nous donne l'image de (x,y) par g :

$$g(x,y)=\tau_{e_1}^k\circ\tau^l_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}\begin{pmatrix}-\frac{1}{2}x-\frac{\sqrt{3}}{2}y\\ -\frac{\sqrt{3}}{2}x-\frac{1}{2}y\end{pmatrix}=\begin{pmatrix}-\frac{1}{2}x-\frac{\sqrt{3}}{2}y+k-\frac{1}{2}l\\ \frac{\sqrt{3}}{2}x-\frac{1}{2}y+\frac{\sqrt{3}}{2}l\end{pmatrix}. \tag{19.526}$$

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 (19.513) mais avec des inégalités strictes. Le fait que $g(x,y)\in\text{Int}(K)$ nous donne trois inéquations de plus. Assez rapide calcul :

$$f_{AB}\big(g(x,y)\big)=-2x+k-2l, \tag{19.527a}$$
$$f_{BC}\big(g(x,y)\big)=-x-\sqrt{3}y-k-l+1, \tag{19.527b}$$
$$f_{AC}\big(g(x,y)\big)=\frac{\sqrt{3}}{2}x-\frac{1}{2}y+\frac{\sqrt{3}}{2}l. \tag{19.527c}$$

Au final, notre point (x,y) doit satisfaire le système suivant :

$$\begin{cases}x-\sqrt{3}y>0 & \text{(19.528a)}\\ -x-\sqrt{3}y+1>0 & \text{(19.528b)}\\ y>0 & \text{(19.528c)}\\ -2x+k-2l>0 & \text{(19.528d)}\\ -x-\sqrt{3}y-k-l+1>0 & \text{(19.528e)}\\ \frac{\sqrt{3}}{2}x-\frac{1}{2}y+\frac{\sqrt{3}}{2}l>0. & \text{(19.528f)}\end{cases}$$

Les inéquations 19.528c et 19.528a donnent déjà $x > 0$. De même avec 19.528b nous trouvons $x < 1$. Voila déjà $x \in]0,1[$ qui est directement visible sur le dessin du triangle K.
Vu que $x > 0$, l'inéquation (19.528d) donne

$$k - 2l > -2x + k - 2l > 0. \tag{19.529}$$

Donc $k - 2l > 0$.
Vu que $x < 1$ et $y > 0$, l'inéquation (19.528f) donne

$$\frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2}l > \frac{\sqrt{3}}{2}x - \frac{1}{2}y + \frac{\sqrt{3}}{2}l > 0. \tag{19.530}$$

Donc $\frac{\sqrt{3}}{2}(1 + l) > 0$. Comme l est entier, cela donne $l \geqslant 0$.
Enfin de (19.528e) nous tirons

$$-k + 1 > -x - \sqrt{3}y - k - l + 1 > 0, \tag{19.531}$$

Ce qui donne $k < 1$ et donc $k \leqslant 0$.
En résumé nous avons trouvé trois inéquations pour k et l :

$$\begin{cases} k \leqslant 0 & \text{(19.532a)} \\ l \geqslant 0 & \text{(19.532b)} \\ k - 2l > 0. & \text{(19.532c)} \end{cases}$$

Ce système est impossible.
Il n'existe donc pas de translations qui, appliquée à $\text{Int}(L)$, donne une intersection avec $\text{Int}(K)$.

Pour $L = R_0(\pi/3)^3K$ hop.

Pour $L = R_0(\pi/3)^4K$ hop.

Pour $L = R_0(\pi/3)^5K$ hop.

Voila. 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 [41]. □

Lemme 19.178.
Le groupe

$$\text{gr}\left(\tau_1, \tau_{(-\frac{1}{2},\frac{\sqrt{3}}{2})}, R_0(\pi/3)\right) \tag{19.533}$$

est un groupe de pavage.

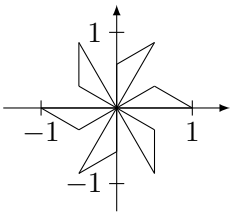
D'après [327], la démonstration du lemme 19.178 demande d'utiliser le losange de sommets $(0,0)$, $(1,0)$, $(\frac{1}{2}, \frac{\sqrt{3}}{6})$ et $(\frac{1}{2}, -\frac{\sqrt{3}}{6})$ [42].

Lemme 19.179 ([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 $\|u_1\| \geqslant \|u_2\|$. Alors

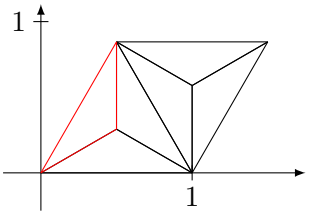
$$\min\{\|v\|, \|u_1 + u_2 - v\|\} \leqslant \|u_1\|. \tag{19.534}$$

Autrement dit, tout point intérieur d'un parallélogramme est plus proche d'un angle que la longueur du plus long côté.

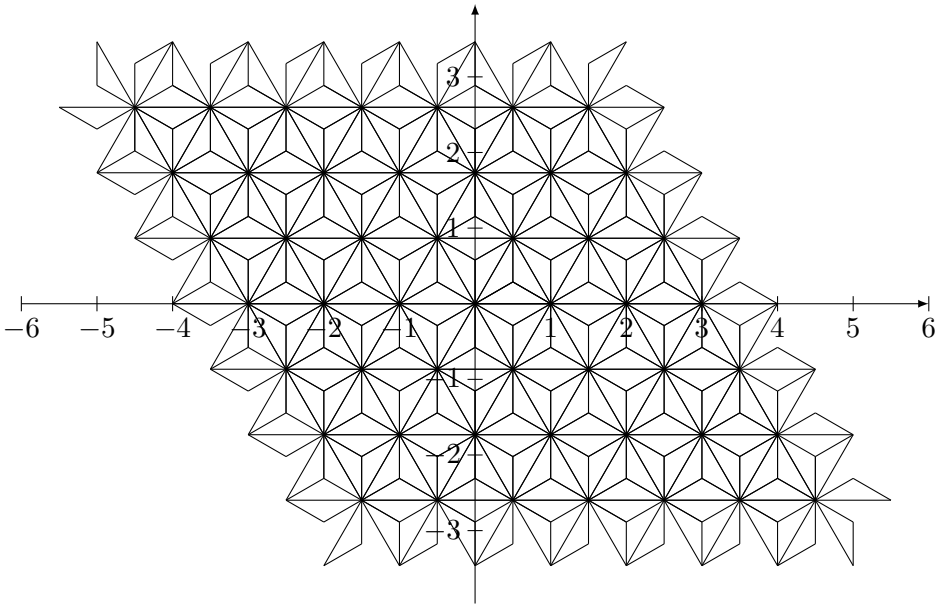
41. Je n'ai pas vérifié. Faites-le et écrivez-moi pour me dire ce qui en est.
42. Je n'ai pas vérifié, mais à mon avis une preuve doit prendre les mêmes idées que celles du lemme 19.177.



(a) Le compact K est ses rotations.



(b) Une maille du réseau des translations de G.



(c) Une partie du pavage complet.

FIGURE 19.10 – Illustrations pour le pavage du lemme 19.177.

Démonstration. Les points $\alpha u_1 + \beta u_2$ ($\alpha, \beta \in [0,1]$) se divisent en deux parties : ceux avec $0 \leqslant \alpha + \beta \leqslant 1$ et ceux avec $1 \leqslant \alpha + \beta \leqslant 2$.

Si $\alpha + \beta \leqslant 1$ alors

$$\|\alpha u_1 + \beta u_2\| < \|\alpha u_1\| + \beta\|u_2\| = \alpha\|u_1\| + \beta\|u_2\| \leqslant (\alpha+\beta)\|u_1\| \leqslant \|u_1\|. \tag{19.535}$$

L'inégalité est stricte parce que u_1 et u_2 ne sont pas colinéaires.

Si au contraire $\alpha + \beta \geqslant 1$ nous avons

$$\|u_1 + u_2 - \alpha u_1 - \beta u_2\| < \|(1-\alpha)u_1\| + \|(1-\beta)u_2\| < (2-\alpha-\beta)\|u_1\| \leqslant \|u_1\|. \tag{19.536}$$

□

Lemme 19.180.
Soient un ensemble E, et deux bijections $r, s \colon E \to E$ 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

$$(rs)(x) = r(a) = a. \tag{19.537}$$

Et d'autre part,

$$(sr)(a) = (srs^{-1})(a) \tag{19.538}$$

Si nous imposons $(srs^{-1})(a) = a$, nous avons, en appliquant s^{-1} est 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 $a = b$. □

Lemme 19.181.
Soit un sous groupe fini 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

$$T=\mathbb{Z}u_1+\mathbb{Z}u_2. \tag{19.539}$$

Démonstration. Nous décomposons en plusieurs parties.

$\mathbb{R}^+v\cap T=\mathbb{N}v_m$ 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 λ_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, doit $\lambda v\in T$. Nous devons prouver que $\lambda=l\lambda_m$ pour un certain $l\in\mathbb{N}$. Soit $k\in\mathbb{N}$ tel que $k\leqslant\lambda<k+1$.

Nous avons

$$(k+1)\lambda_m-\lambda\leqslant(k+1)\lambda_m-k\lambda_m=\lambda_m. \tag{19.540}$$

Vu que T est un groupe pour l'addition, les faits que $\lambda_m v\in T$ et que $\lambda v\in T$ impliquent que $\big((k+1)\lambda_m-\lambda\big)v\in T$. Mais

$$|\big((k+1)\lambda_m-\lambda\big)v|\leqslant\lambda_m. \tag{19.541}$$

Vu la propriété de minimalité de λ_m nous avons forcément

$$(k+1)\lambda_m-\lambda=\lambda_m. \tag{19.542}$$

Cela prouve que $\lambda=k\lambda_m$.

Jusqu'ici nous avons prouvé que

$$\mathbb{R}^+v\cap T=\mathbb{N}v_m \tag{19.543}$$

pour un certain multiple v_m de v.

$\|u_1\|\leqslant\|v\|$ **pour tout** v Nous montrons à présent qu'il existe $u_1\in T$ tel que $\|u_1\|\leqslant\|v\|$ 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 le couronne (compacte) de rayons δ et $\|u_1\|$. Quitte à prendre une sous-suite, nous pouvons supposer que (v_k) converge[43]. Cette suite serait de Cauchy et pour tout ϵ (en particulier $\epsilon<\delta$), il existeraient p,q tels que $\|v_p-v_q\|<\epsilon$. Vu que T est un groupe pour l'addition, nous aurions $v_p-v_q\in T$ avec $\|v_p-v_q\|<\epsilon\leqslant\delta$.

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.

$T=\mathbb{Z}u_1+\mathbb{Z}u_2$ 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.

Soit $v\in T$. Vu que $\{u_1,u_2\}$ est une base de $\mathbb{R}^2$, il existe $\alpha,\beta\in\mathbb{R}$ tels que

$$v=\alpha u_1+\beta u_2. \tag{19.544}$$

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 19.179. 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.

□

43. Proposition 8.19.

Proposition 19.182 ([1]).
Si G est un groupe de pavage[44] *de $\mathbb{R}^2$ et si $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ est une isométrie affine, alors le groupe*

$$G' = f \circ G \circ f^{-1} = \{f \circ g \circ f^{-1} \text{ tel que } g \in G\} \tag{19.545}$$

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 fait trois éléments à prouver.

<u>G' **est constitué de déplacements**</u> Les éléments de G sont des isométries, ainsi que f et f^{-1}. Donc les éléments de G' sont des isométries.

Soit $g \in G$. Vu que 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 10.56(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 11.52 qui nous indique que

$$\det(f_0 g_0 f_0^{-1}) = \det(f_0)\det(g_0)\det(f_0^{-1}) = \det(f_0)\det(g_0)\det(f_0)^{-1} = \det(g_0) = 1. \tag{19.546}$$

<u>$G' \cdot K' = \mathbb{R}^2$</u> Vu que f est affine, K' est encore compact[45] Nous avons :

$$\begin{aligned}
(f \circ G \circ f^{-1})\big(f(K)\big) &= \bigcup_{g\in G}\bigcup_{k\in K}(f \circ g \circ f^{-1})\big(\alpha(k)\big) && \text{(19.547a)}\\
&= \bigcup_{g\in G}\bigcup_{k\in K}(f \circ g)(k) && \text{(19.547b)}\\
&= \bigcup_{g\in G} f\Big(\bigcup_{k\in K} g(k)\Big) && \text{(19.547c)}\\
&= \bigcup_{g\in G}(f \circ g)(K) && \text{(19.547d)}\\
&= f\Big(\bigcup_{g\in G} g(K)\Big) && \text{(19.547e)}\\
&= f\big(G \cdot K\big) && \text{(19.547f)}\\
&= \mathbb{R}^2. && \text{(19.547g)}
\end{aligned}$$

Pour la dernière égalité, nous avons utilisé le fait que f est bijective et que $G \cdot K = \mathbb{R}^2$.

<u>**L'autre condition**</u> Deux éléments de G' s'écrivent fg_1f^{-1} et fg_2f^{-1}. Nous les supposons tels que

$$(fg_1f^{-1}) \cdot \operatorname{Int}(K') \cap (fg_2f^{-1}) \cdot \operatorname{Int}(K') \neq \varnothing. \tag{19.548}$$

Nous avons :

$$\begin{aligned}
(fg_1f^{-1}) \cdot \operatorname{Int}(K') \cap (fg_2f^{-1}) \cdot \operatorname{Int}(K') &= f\big(g_1 \cdot \operatorname{Int}(K)\big) \cap f\big(g_2 \cdot \operatorname{Int}(K)\big) && \text{(19.549a)}\\
&= f\big(\underbrace{g_1 \cdot \operatorname{Int}(K) \cap g_2 \cdot \operatorname{Int}(K)}_{\neq\varnothing}\big) && \text{(19.549b)}\\
&\neq \varnothing. && \text{(19.549c)}
\end{aligned}$$

Justifications :

44. Définition 19.172.
45. Il suffit de prouver que $f(K)$ est fermé et borné par le théorème de Borel-Lebesgue 8.9.

— Égalité (19.549a) parce qu'un peu de topologie nous enseigne que $f^{-1}\big(\text{Int}(K')\big) = \text{Int}\big(f^{-1}(K')\big) = \text{Int}(K)$ parce que f est affine.

— Égalité (19.549b) parce que, f étant bijective, $f(A) \cap f(B) = f(A \cap B)$;

□

Théorème 19.183 ([234, 327])**.**
Nous notons τ_v la translation de vecteur v, $r_{A,\theta}$ la rotation de centre A et d'angle θ 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 $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ telle que $f \circ G \circ f^{-1}$ est un groupe de la liste suivante :

(1) $\text{gr}\big(\tau_1, \tau_2\big)$

(2) $\text{gr}\big(\tau_1, \tau_2, R_0(\pi)\big)$

(3) $\text{gr}\big(\tau_1, \tau_{(\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(2\pi/3)\big)$

(4) $\text{gr}\big(\tau_1, \tau_2, R_0(\pi/2)\big)$

(5) $\text{gr}\big(\tau_1, \tau_{(-\frac{1}{2}, \frac{\sqrt{3}}{2})}, R_0(\pi/3)\big)$.

Démonstration. Les lemmes 19.174, 19.175, 19.176, 19.177, et 19.178 montrent que les groupes listés sont des groupes de pavage. La proposition 19.182 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 α 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,

$$T = \{v \in \mathbb{R}^2 \text{ tel que } \tau_v \in G\}. \tag{19.550}$$

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

$$m \in \tau_v\big(\text{Int}(K)\big) \cap \text{Int}(K) \tag{19.551}$$

par hypothèse, cette intersection est non vide seulement si $v = 0$.

Donc il existe δ tel que $\|v\| \geqslant \delta$ pour tout $v \in T$.

Utilisation du lemme La partie T est donc dans la position du lemme 19.181 et nous avons

$$T = \mathbb{Z}u_1 + \mathbb{Z}u_2 \tag{19.552}$$

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 translations. 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 $u_1, u_2 \in \mathbb{R}^2 \backslash \{0\}$.

Translation Si $r, s \in G$, alors l'élément $rsr^{-1}s^{-1}$ est une translation. En effet, vu que 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 τ_r, τ_s telles que $r = A_r \circ \tau_r$ et $s = A_s \circ \tau_s$. Le lemme 10.56 nous donne les inverses. Nous avons

$$rsr^{-1}s^{-1} = (A_r \circ \tau_r)(A_s \circ \tau_s)(A_r^{-1} \circ \tau_{-A_r v_r})(A_s^{-1} \circ \tau_{-A_s v_s}). \tag{19.553}$$

La partie linéaire de cela est

$$A_r \circ A_s \circ A_r^{-1} \circ A_s^{-1}. \tag{19.554}$$

C'est donc une composée de rotations centrées en $(0,0)$. Mais ces rotations forment un groupe abélien (proposition 19.128). Donc nous pouvons écrire

$$A_r \circ A_s \circ A_r^{-1} \circ A_s^{-1} = A_r \circ A_r^{-1} \circ A_s \circ A_s^{-1} = \mathrm{Id}\,. \tag{19.555}$$

Tout ceci pour dire que dès que $r, s \in G$, l'élément $rsr^{-1}s^{-1}$ est une translation.

Les parties linéaires[1] Nous savons de l'exemple 10.13 que les éléments de G s'écrivent sous la forme $f = \tau_v \circ \alpha$ où $v \in \mathbb{R}^2$ et $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ est linéaire.

De plus, f étant une isométrie de $(\mathbb{R}^2, d)$, l'application α est une isométrie de $(\mathbb{R}^2, \|.\|)$. Vu que α est une isométrie, $\det(\alpha) = \pm 1$. Mais les déplacement conservent l'orientation ; donc α doit conserver l'orientation, et la proposition 11.69 nous dit que $\det(\alpha) > 0$. Donc

$$\det(\alpha) = 1. \tag{19.556}$$

Le théorème 19.83 dit que α est la composition d'un nombre pair de réflexions. Mais comme il y en a au plus trois (théorème 19.161), l'application α est composée de zéro ou deux réflexions.

Donc les parties linéaires des éléments de G sont des rotations.

Les autres Les parties linéaires des éléments de G sont des rotations. Mais les éléments de G eux-mêmes ne sont pas tellement mystérieux. Vu que ce sont des isométries de $(\mathbb{R}^2, d)$, elles sont composées de 0, 1, 2 ou 3 réflexions.

Mais ce sont des déplacement, donc ils préservent l'orientation et le théorème 19.83 dit qu'ils sont des composées de zéro ou deux réflexions (nombre pair). Ce sont donc des rotations.

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 préciserons « 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 [46].

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 α 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

$$(\alpha\tau_v\alpha^{-1})(x) = \alpha\big(\alpha^{-1}(x) + v\big) = x + \alpha(v) = \tau_{\alpha(v)}(x). \tag{19.557}$$

Donc $\alpha\tau\alpha^{-1} = \tau_{\alpha(v)}$ et $\alpha(v) \in T$.

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. Vu qu'une rotation a son centre comme unique point fixe, le lemme 19.180 nous dit que toutes le 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 $K \subset B(c, r)$. Vu que G stabilise toutes les boules centrées en c, nous avons

$$G \cdot K \subset B(c, r). \tag{19.558}$$

Donc nous n'avons pas un recouvrement de $\mathbb{R}^2$. Le cas $T = \{0\}$ est exclu.

Exclusion de $T = \mathbb{Z}u$ Nous supposons à présent que $T = \mathbb{Z}u$ pour un certain $u \in \mathbb{R}^2$.

46. 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.

$r_0 = \pm \mathrm{Id}$ Nous savons que tous les éléments de G sont des rotations ; soit un élément r de G. La proposition 19.127(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 $\pm \mathrm{Id}$. D'abord,

$$r\tau_u r^{-1} = \tau_a^{-1} r_0 \tau_a \tau_u \tau_a^{-1} r_0^{-1} \tau_a = \tau_a^{-1} r_0 \tau_u r_0^{-1} \tau_a. \tag{19.559}$$

Ensuite, nous appliquons cela à $x \in \mathbb{R}^2$:

$$\begin{aligned}(\tau_a^{-1} r_0 \tau_u r_0^{-1} \tau_a)(x) &= (\tau_a^{-1} r_0 \tau_u)\big(r_0^{-1}(x+a)\big) && (19.560a)\\ &= (\tau_a^{-1} r_0)\big(r_0^{-1}(x+a) + u\big) && (19.560b)\\ &= \tau_a^{-1}\big(x + a + r_0(u)\big) && (19.560c)\\ &= x + r_0(u). && (19.560d)\end{aligned}$$

Donc $r \circ \tau_u \circ r^{-1} = \tau_{r_0(u)}$, ce qui prouve que $r_0(u) \in T$. Vu que $\|r_0(u)\| = \|u\|$ nous avons forcément $r_0(u) = \pm u$.

Si $r_0(u) = u$, alors $r_0 = \mathrm{Id}$ (parce que r_0 est une rotation fixant plus que un seul point). Dans ce cas, $r = \mathrm{Id}$.

Si au contraire $r_0(u) = -u$, alors $r_0 = -\mathrm{Id}$.

Forme générale Donc si r est un élément non trivial de G nous avons $r = \tau_a^{-1} \circ (-\mathrm{Id}) \circ \tau_a$, et alors

$$\big(\tau_a^{-1} \circ (-\mathrm{Id}) \circ \tau_a\big)(x) = \tau_a^{-1}(-\mathrm{Id})(x+a) = \tau_a^{-1}(-x-a) = -x - 2a. \tag{19.561}$$

Donc pour tout $r \in G$, il existe $a \in \mathbb{R}^2$ tel que

$$r(x) = -x - 2a. \tag{19.562}$$

Pour information, le centre de cette rotation est $-a$ (c'est le seul point fixe).

Les centres sont alignés Soit r une rotation de centre $-a$ et s de centre $-b$. Alors

$$(rs)(x) = r(-x-2b) = x + 2b - 2a = x + 2(b-a). \tag{19.563}$$

Donc $rs = \tau_{2(b-a)}$.

Cela prouve que $2(b-a) \in \mathbb{Z}u$.

Une bande Soit la droite $D = \mathbb{R}u$. Nous considérons la bande

$$B_r = \{x \in \mathbb{R}^2 \text{ tel que } d(x, D) < r\}. \tag{19.564}$$

Une inclusion Nous prouvons à présent que pour tout r, nous avons

$$G \cdot B_r \subset B_r. \tag{19.565}$$

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 (19.562). Nous avons

$$\begin{aligned}d\big(r(x), D\big) &= d(x + 2a, D) && (19.566a)\\ &= d\big(\tau_{2a}^{-1}(x+2a), \tau_{2a}^{-1}(D)\big) && (19.566b)\\ &= d(x, D). && (19.566c)\end{aligned}$$

Justifications :

- Pour (19.566a), nous avons $D = -D$ et $d(x,y) = d(-x,-y)$.
- Pour (19.566b), invariance par translation de la distance dans $\mathbb{R}^2$.
- Pour (19.566c), les éléments de D sont les multiples de a ; donc cette droite est invariante par cette translation.

Bref, si $d(x,D)=d\big(r(x),D\big)$. Donc pour tout $r>0$[47] et pour tout $g\in G$, si $x\in B_r$, alors $g(x)\in B_r$.

Exclusion Vu que 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 par tout $\mathbb{R}^2$ et G n'est pas un groupe de pavage.

Et nous voila avec seulement $T=\mathbb{Z}u_1+\mathbb{Z}u_2$ en lice.

Pause : quelque parties de G à ne pas confondre Il y a plusieurs parties de G qu'il convient de ne pas confondre. Je vous laisse méditer quelque temps sur la liste suivante :

(1) G est le groupe que nous cherchons à déterminer ;

(2) T est le groupe des translations de G ;

(3) le groupe des rotations linéaires dans G ;

(4) l'ensemble des $\tau_A^{-1}\circ r\circ\tau_A$ où r est une rotation de G centrée en A.

(5) L est l'ensemble des parties linéaires des éléments de G.

En particulier les deux derniers ne sont pas les mêmes.

Dans le cas (4), 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 19.127(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 garanties que cela forme un groupe.

Si r est une rotation dans G, dans le cas (5) il s'agit de décomposer $r=\tau_v\circ\alpha$ (lemme 10.9) et considérer α. Dans ce cas, α est linéaire, mais il n'y a pas de garanties que α soit une rotation.

L est un groupe En trois conditions.

— 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 10.56(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 liénaire d'un élément de G.

— De la même façon, en utilisant l'inverse 10.56(3), $(\tau_v\circ\alpha)^{-1}=\tau_{-\alpha^{-1}(v)}\circ\alpha^{-1}$. Donc α^{-1} est la partie linéaire d'un élément de G.

— Et enfin $\mathrm{Id}=\tau_0\circ\mathrm{Id}$. Donc l'identité est dans L.

Ok pour que L soit un groupe.

Précision L'ensemble L est un groupe certes. Mais rien ne dit que L soit un sous-groupe de G.

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 τ_u sont dans G, l'élément $g\tau_u g^{-1}$ est également dans G. Nous l'appliqons à $x\in\mathbb{R}^2$:

$$\begin{aligned}(\tau_v\alpha\tau_u\alpha^{-1}\tau_v^{-1})(x)&=(\tau_v\alpha\tau_u\alpha^{-1})\alpha^{-1}(x-v)&(19.567a)\\&=(\tau_v\alpha)\big(\alpha^{-1}(x-v)+u\big)&(19.567b)\\&=\tau_v\big(x-v+\alpha(u)\big)&(19.567c)\\&=x+\alpha(u)&(19.567d)\\&=\tau_{\alpha(u)}(x).&(19.567e)\end{aligned}$$

Donc $g\tau_u g^{-1}=\tau_{\alpha(u)}\in G$.

Question de trace Soit $\alpha\in L$; dans la base $\{u_1,u_2\}$ la matrice de α est

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\tag{19.568}$$

47. Remarquez la notation malheureuse pour r qui est maintenant une distance alors que trois mots plus tôt, c'était un élément de G.

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 19.138 nous dit qu'il existe $\theta \in [0, 2\pi[$ tel que la matrice de α soit

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{19.569}$$

La trace est $2\cos(\theta)$. La trace est invariante par changement de base[48], donc $2\cos(\theta) = a + d \in \mathbb{Z}$. Les possibilités pour $\cos(\theta)$ sont donc -1, $-1/2$, 0, $1/2$ et 1.

Les angles possibles Nous savons que $\cos(\frac{\pi}{2}) = 0$ et $\cos(\pi/3) = 1/2$ (proposition 19.32(6) et lemme 19.35). La proposition 19.68 nous dit alors que, dans notre cas, les valeurs possibles pour θ dans $[0, 2\pi[$ sont

$$\{0, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \pi, \frac{4\pi}{3}, \frac{3\pi}{2}, \frac{5\pi}{3}\}. \tag{19.570}$$

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.

Quelque combinaisons impossibles Vu que 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[49] $R_0(\pi/2)R_0(\pi/3) = R_0(5\pi/6)$ est également dans L. Mais comme $5\pi/6$ n'est pas dans la liste (19.570), $R_0(5\pi/6)$ n'est pas dans L.

En raisonnant de la sorte, nous voyons que si $R_0(\pi/2) \in L$, alors $L = \text{gr}\big(R_0(\pi/2)\big)$.

La liste Plus généralement, les possibilités pour L sont

- $\{\text{Id}\}$
- $\text{gr}\big(R_0(\pi/2)\big)$
- $\text{gr}\big(R_0(\pi/3)\big)$
- $\text{gr}\big(R_0(\pi)\big)$
- $\text{gr}\big(R_0(2\pi/3)\big)$

Une justification plus courte pour cette liste est d'invoquer le théorème 19.164(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 $\text{gr}\big(R_0(4\pi/3)\big) = \text{gr}\big(R_0(2\pi/3)\big)$ 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 $R_0(5\pi/3) = R_0(\pi/3)^{-1}$.

Le cas un peu générique Nous supposons que $L = \text{gr}\big(R_0(\theta)\big)$ pour un certain θ. Nous allons voir qu'à part dans les cas $\theta = 0$ et $\theta = \pi$, il est possible de trouver une application affine α telle que le groupe $\alpha G \alpha^{-1}$ est 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[50] $G' = \tau_{w_0} G \tau_{w_0}^{-1}$ 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 $x \in \mathbb{R}^2$:

$$g(x) = R_0(\theta)x - R_0(\theta)w_0 + w_0 + v_0. \tag{19.571}$$

Nous avons $g(x) = R_0(\theta)x$ lorsque

$$w_0 = \big(R_0(\theta) - \mathbb{1}\big)^{-1} v_0. \tag{19.572}$$

Voila 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 $R_0(\theta)$. Vu qu'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

$$\tau_{w_0} \circ \tau_v \circ R_0(\theta)^k \circ \tau_{w_0} = \tau_{w_0 + v - R_0(\theta)^k w} \circ R_0(\theta)^k. \tag{19.573}$$

48. Proposition 11.209.
49. Proposition 19.139 pour l'addition des angles.
50. Proposition 19.182.

Donc tous les éléments de G' sont encore de la forme

$$\tau_v \circ R_0(\theta)^k. \tag{19.574}$$

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

$$G' = \text{gr}\big(\tau_{u_1}, \tau_{u_2}, R_0(\theta)\big). \tag{19.575}$$

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 $u_2 = R_0(\theta)$ 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

$$G' = \text{gr}\big(\tau_{u_1}, \tau_{R_0(\theta)u_1}, R_0(\theta)\big). \tag{19.576}$$

Nous allons maintenant nous occuper de u_1. Pour cela nous considérons une rotation suivie d'une homothétie α telle que $\alpha(u_1) = e_1$. Une telle opération α d'une part commute avec $R_0(\theta)$ et d'autre part fait $\alpha \circ \tau_v \circ \alpha^{-1} = \tau_{\alpha(v)}$. Donc le groupe $G'' = \alpha G' \alpha^{-1}$ est, par le lemme 2.9,

$$G'' = \text{gr}\big(\tau_{e_1}, \tau_{\alpha R_0(\theta)u_1}, R_0(\theta)\big). \tag{19.577}$$

Cela s'écrit aussi bien

$$G'' = \text{gr}\big(\tau_{e_1}, \tau_{R_0(\theta)e_1}, R_0(\theta)\big). \tag{19.578}$$

Et voila, ce groupe G'' est un de la liste.

Le cas $L = \{\text{Id}\}$ Dans ce cas, tous les éléments de G sont de la forme τ_v avec $v \in T$. Nous considérons l'application linéaire $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ telle que $\alpha(u_1) = e_1$ et $\alpha(u_2) = e_2$. Nous avons

$$(\alpha \circ \tau_{u_i} \circ \alpha^{-1})(x) = (\alpha\tau_{u_i})\alpha^{-1}(x) = \alpha\big(\alpha^{-1}(x) + u_i\big) = x + \alpha(u_i) = \tau_i(x). \tag{19.579}$$

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

$$\alpha\tau_v\alpha^{-1} = \alpha\tau_{au_1}\tau_{au_2}\alpha^{-1} = \alpha\tau_{au_1}\alpha^{-1}\alpha\tau_{bu_2}\alpha^{-1} = \tau_{ae1}\tau_{be_2}. \tag{19.580}$$

Nous avons donc

$$\alpha G\alpha^{-1} = \text{gr}(\tau_1, \tau_2). \tag{19.581}$$

Cela est un des groupes de la liste.

Notez qu'à la place de ces calculs, nous pouvions aussi invoquer la proposition 2.9.

Le cas $L = \{-\text{Id}\}$ Dans le cas $\theta = \pi$, nous pouvons aller jusqu'à (19.575) et écrire

$$G' = \text{gr}\big(\tau_{u_1}, \tau_{u_2}, R_0(\pi)\big). \tag{19.582}$$

Vu que $R_0(\pi) = -\text{Id}$ commute avec toutes les applications linéaires et que u_1 et u_2 ne sont pas colinéaires, nous pouvons considérer une application linéaire $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$ telle que $\alpha(u_1) = e_1$, $\alpha(u_2) = e_2$. Nous avons alors

$$G'' = \alpha G'\alpha^{-1} = \text{gr}\big(\tau_{e_1}, \tau_{e_2}, -\text{Id}\big), \tag{19.583}$$

qui est encore dans la liste.

□

19.11 Un peu de structure de $\mathrm{O}(n)$

19.11.1 Valeurs propres dans $\mathrm{O}(n)$

Proposition 19.184 ([328]).
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 $Ax = \lambda x$. Vu que les éléments de la matrice A sont réels,

$$A\bar{x} = \bar{A}\bar{x} = \overline{Ax} = \overline{\lambda x} = \bar{\lambda}\bar{x}. \tag{19.584}$$

Donc $\bar{\lambda}$ est une valeur propre de A pour le vecteur propre $\bar{x}$.

Soit λ une valeur propre de A de vecteur propre x. Alors nous avons d'une part

$$\langle \overline{Ax}, Ax \rangle = \langle A^t A\bar{x}, x \rangle = \langle x, \bar{x} \rangle, \tag{19.585}$$

et d'autre part

$$\langle \overline{Ax}, Ax \rangle = \langle \bar{\lambda}\bar{x}, \lambda x \rangle = |\lambda|^2 \langle \bar{x}, x \rangle. \tag{19.586}$$

Vu que $x \neq 0$ nous avons aussi $\langle \bar{x}, x \rangle \neq 0$. Par conséquent $|\lambda|^2 = 1$ et $|\lambda| = 1$. □

Lemme 19.185 ([328]).
Soit 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 $f_F \colon F \to F$ 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 :

$$\langle y, f(a) \rangle = \langle f(x), f(a) \rangle = \langle x, a, \rangle = 0 \tag{19.587}$$

parce que $x \in F^\perp$. □

Proposition 19.186 ([328]).
Soit une isométrie $f \colon \mathbb{R}^3 \to \mathbb{R}^3$.

(1) L'application linéaire f possède au moins une valeur propre réelle qui vaut ± 1.

(2) Il existe une base orthonormée de $\mathbb{R}^3$ dans laquelle la matrice de f est de la forme

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \cos(\theta) & -\epsilon\cos(\theta) \\ 0 & \sin(\theta) & \epsilon\cos(\theta) \end{pmatrix} \tag{19.588}$$

avec $\epsilon, \lambda = \pm 1$ et $\theta \in [0, 2\pi[$.

Démonstration. Le polynôme caractéristique de f, donné par $\det(f - \lambda\,\mathrm{Id})$, est à coefficients réels et de degré 3. Il possède dont au moins une solution réelle par le corollaire 13.99. Soit donc une valeur propre réelle λ de χ_f ; par le lemme 19.184 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 19.185. Soit une base orthonormée $\{u_2, u_3\}$ de $F^\perp$ et la matrice B de la restriction f_p à $F^\perp$. Vu que l'application f_p est une isométrie de $F^\perp$, la matrice B est, par le lemme 19.134, de la forme

$$B = \begin{pmatrix} \cos(\theta) & -\epsilon\sin(\theta) \\ \sin(\theta) & \epsilon\cos(\theta) \end{pmatrix} \tag{19.589}$$

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

$$\begin{pmatrix} \lambda & 0 \\ 0 & B \end{pmatrix}, \tag{19.590}$$

comme annoncé. □

Pour classifier les isométries de $\mathbb{R}^3$, nous pouvons nous baser sur les possibilités de la matrice donne dans le lemme 19.186. Il y a essentiellement quatre possibilités suivant les valeurs de $\lambda = \pm 1$ et $\epsilon = \pm 1$.

Si $\epsilon = \lambda = 1$ Alors la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix} \tag{19.591}$$

et l'isométrie correspondante est la rotation d'angle $-\theta$ autour de la droite de u_1.

Si $\epsilon = \lambda = -1$ Alors la matrice est

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{pmatrix} \tag{19.592}$$

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 $\{u_1, u_2, u_3\}$, les points de cette droite sont de la forme [51]

$$\big(0, \cos(\theta/2), \sin(\theta/2)\big). \tag{19.593}$$

L'image de u_1 par cette réflexion est $-u_1$, c'est clair.

Faisons 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.

Dans le plan $\text{Span}(u_2, u_3)$ 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

$$M = \left(\frac{\sin(\theta)}{2}, \frac{1 - \cos(\theta)}{2} \right). \tag{19.594}$$

Les formules de duplication d'angle du corollaire 19.27 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 :

$$\begin{aligned} M &= \left(\cos(\theta/2)\sin(\theta/2), \frac{1 - \big(\cos^2(\theta/2) - \sin^2(\theta/2)\big)}{2} \right) && \text{(19.595a)} \\ &= \big(\cos(\theta/2)\sin(\theta/2), \sin^2(\theta/2) \big) && \text{(19.595b)} \\ &= \sin(\theta/2)\big(\cos(\theta/2), \sin(\theta/2) \big). && \text{(19.595c)} \end{aligned}$$

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

$$\big(A(u_3) - u_3\big) \cdot \begin{pmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = \sin(\theta)\cos(\theta/2) - \big(1 + \cos(\theta)\big)\sin(\theta/2) = 0 \tag{19.596}$$

où nous avons encore utilisé les duplications d'angles et le fait que $1 = \cos^2(\theta/2) + \sin^2(\theta/2)$ (lemme 19.18).

51. Les plus acharnés remarquerons que $\{u_1, u_2, u_3\}$ est un ensemble, qui est une base. Mais un ensemble n'est pas ordonné, alors que pour écrire l'équation de droite qui suit, nous supposons un ordre. Je laisse au tel lecteur le soin de trouver une bonne notation.

Si $\epsilon = -1$ et $\lambda = 1$ Alors la matrice est

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & \sin(\theta) & -\cos(\theta) \end{pmatrix}. \tag{19.597}$$

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,

$$A(v) = \cos(\theta/2)\begin{pmatrix} 0 \\ \cos(\theta) \\ \sin(\theta) \end{pmatrix} + \sin(\theta/2)\begin{pmatrix} 0 \\ -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = \begin{pmatrix} 0 \\ \cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} = v. \tag{19.598}$$

Nous avons sauté quelques étapes de calcul mettant en scène les formules de duplication d'angle : 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 :

$$\big(\cos(\theta/2)u_2 + \sin(\theta/2)u_3\big) \cdot (xu_1 + yu_2 + zu_3) = 0 \tag{19.599a}$$

$$u_1 \cdot (xu_1 + yu_2 + zu_3) = 0. \tag{19.599b}$$

Nous avons immédiatement $x = 0$ et ensuite la relation

$$y\cos(\theta/2) + z\sin(\theta/2) = 0. \tag{19.600}$$

En ne regardant que les deux dernières composantes pour alléger l'écriture,

$$A(w) = y\begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} + z\begin{pmatrix} \sin(\theta) \\ -\cos(\theta) \end{pmatrix}. \tag{19.601}$$

Le but est de montrer que cela est égal à $-y\cos(\theta/2) - z\sin(\theta/2)$.

Notons $c = \cos(\theta/2)$ et $s = \sin(\theta/2)$. Alors $A(w)_2 = y(c^2 - s^2) + 2zcs$. Évacuons tout de suite les deux cas limite : si $c = 0$ alors $A(w)_2 = -y$ (parce que $s = \pm 1$) et c'est bon. Si $s = 0$, alors $A(w)_2 = y$, mais la relation (19.600) donne $y = 0$, donc c'est bon aussi. Dans la cas générique, $z = -yc/2$ et

$$A(w)_2 = y(c^2 - s^2) - 2cs\frac{yc}{s} = -y(c^2 + s^2) = -y. \tag{19.602}$$

En ce qui concerne $A(w)_3$, c'est très similaire :

$$A(w)_3 = 2ysc - z(c^2 - s^2). \tag{19.603}$$

Avec $z = 0$ c'est $-z$, donc c'est bon. Avec $c = 0$ c'est z mais $z = 0$. Et pour le cas générique, la substitution $y = -zs/c$ donne le résultat.

Si $\epsilon = 1$ et $\lambda = -1$ Alors la matrice est

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{19.604}$$

Cela est la composition entre la symétrie de plan $\mathrm{Span}(u_2, u_3)$ et la rotation d'angle θ dans ce plan.

19.11.2 Sous-groupes finis de SO(3)

Lemme 19.187 ([1]).
Points fixes pour SO(3).

(1) *Tout élément de* SO(3) *possède une droite de point 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 dont (en comptant les multiplicités) trois racines dont une réelle par le corollaire 13.99. Vu que nous sommes en dimension impaire, le coefficient du terme de degré 3 est -1 et le polynôme caractéristique de $g \in \mathrm{SO}(3)$ s'écrit

$$\chi_g(X) = -(X - \lambda_1)(X - \bar{\lambda}_1)(X - s) \tag{19.605}$$

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

$$1 = \det(g) = \chi_g(0) = \lambda_1 \bar{\lambda}_1 s = s. \tag{19.606}$$

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.

Vu que g est une isométrie, la base reste orthonormée sous l'action de g. Donc g a pour matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}. \tag{19.607}$$

Pour que le déterminant soit 1, il faut que la matrice soit l'identité. □

Proposition 19.188 ([1, 329, 57, 330, 276]).
Les sous-groupes finis de SO(3) *sont :*

(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 groups symétrique* S_4.

Démonstration. Soit G, un sous-groupe fini de SO(3). Par la proposition 11.83, les éléments de G sont des isométries de $\mathbb{R}^3$, et le lemme 19.187 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 Ω l'ensemble des pôles des éléments non triviaux de G.

Une action Le groupe G agit sur Ω. 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. Par dur : $(hgh^{-1})h(x) = h(x)$; donc $h(x)$ est un pôle de hgh^{-1}.

Les fixateurs sont cycliques Nous montrons à présent que pour tout $x \in \Omega$, le sous-groupe $\mathrm{Fix}(x)$ est cyclique. Soit donc $x \in \Omega$, le plan orthogonal $\sigma = \mathrm{Span}(x)^\perp$ et $h \in \mathrm{Fix}(x)$. Nous avons $h(\sigma) = \sigma$. En effet si $y \in \sigma$ nous avons

$$0 = y \cdot x = h(y) \cdot h(x) = h(y) \cdot x, \tag{19.608}$$

donc $h(y)$ est perpendiculaire à x. L'inclusion inverse se démontre de même : si $y \in \sigma$ alors $y = h\big(h^{-1}(y)\big)$ alors que $h^{-1}(y) \in \sigma$.

La restriction de h à σ est une isométrie de σ. Prenant une isométrie $f\colon \sigma \to \mathbb{R}^2$, l'application

$$\begin{aligned}\varphi\colon \operatorname{Fix}(x) &\to \mathrm{SO}(2)\\ h &\mapsto f\circ h\circ f^{-1}.\end{aligned} \tag{19.609}$$

est un morphisme injectif de groupes. En effet nous avons d'une part

$$\varphi(hh') = f\circ h\circ h'\circ f^{-1} = fhf^{-1}fh'f^{-1} = \varphi(h)\varphi(h'), \tag{19.610}$$

d'où le morphisme. Et d'autre part, si $\varphi(h) = \varphi(h')$ alors $f\circ h\circ f^{-1} = g\circ h'\circ f^{-1}$, qui donne immédiatement $h = h'$.

Nous en déduisons que $\operatorname{Fix}(x)$ est isomorphe à un sous-groupe de $SO(2)$ (l'image de φ). Le lemme 19.141 en fait un groupe cyclique.

Taille des fixateurs Soient Ω_i les orbites. Si $x, y \in \Omega_i$ alors nous montrons que $|\operatorname{Fix}(x)| = |\operatorname{Fix}(y)|$ avec la bijection

$$\begin{aligned}\varphi\colon \operatorname{Fix}(x) &\to \operatorname{Fix}(y)\\ h &\mapsto g^{-1}hg\end{aligned} \tag{19.611}$$

où g est choisi de façon à avoir $y = g(x)$ (possible parce que x et y sont dans la même orbite). Cela est surjectif parce que si $k \in \operatorname{Fix}(x)$ alors $k = \varphi(gkg^{-1})$ et l'on vérifie que $gkg^{-1} \in \operatorname{Fix}(y)$. L'application φ est également injective parce que si $ghg^{-1} = gh'g^{-1}$ alors $h = h'$.

Un peu de notations Vu que tous les fixateurs des éléments d'une orbite ont la même taille (finie), nous pouvons noter

$$n_i = |\operatorname{Fix}(x_i)| \tag{19.612}$$

pour $x_i \in \Omega_i$. Nous notons également r le nombre d'orbites de G.

La formule de Burnside du théorème 2.56, avec les notations d'ici, donne

$$r = \frac{1}{|G|}\sum_{g\in G} |\operatorname{Fix}(g)|. \tag{19.613}$$

Une belle formule Soit l'ensemble

$$A = \{(g, x) \text{ tel que } g \in G\backslash\{e\}, x \in \operatorname{Fix}(g)\} \tag{19.614}$$

où par $\operatorname{Fix}(g)$ 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 $|\operatorname{Fix}(x)| - 1$ éléments de $G\backslash\{e\}$ qui fixent x. Donc

$$|A| = \sum_{x\in\Omega} \big(|\operatorname{Fix}(x)| - 1\big). \tag{19.615}$$

Mais $|\operatorname{Fix}(x)|$ est constant sur les orbites. Nous coupons donc la somme sur Ω en plusieurs sommes sur les orbites Ω_i :

$$|A| = \sum_{i=1}^{r}\sum_{x\in\Omega_i} \big(|\operatorname{Fix}(x)| - 1\big) = \sum_i |\Omega_i|(n_i - 1). \tag{19.616}$$

En égalisant les deux façons de calculer $|A|$, nous déduisons la formule

$$2\big(|G| - 1\big) = \sum_i |\Omega_i|(n_i - 1). \tag{19.617}$$

Nous utilisons ensuite la relation orbite-stabilisateur, proposition 2.49 : $|\operatorname{Fix}(x_i)||\Omega_i| = |G|$; la formule (19.617) devient

$$2\big(|G|-1\big) = \sum_i |G| - \sum_i \frac{|G|}{n_i} = r|G| + |G|\sum_i \frac{1}{n_i}, \tag{19.618}$$

ou encore, en simplifiant par $|G|$:

$$2 - \frac{2}{|G|} = r - \sum_i \frac{1}{n_i} = \sum_{i=1}^{r}\left(1 - \frac{1}{n_i}\right). \tag{19.619}$$

Nous pouvons aussi repartir de (19.617) et sommer de façon plus simple $\sum_i |\Omega_i| = |\Omega|$ pour obtenir

$$2\big(|G|-1\big) = r|G| - |\Omega| \tag{19.620}$$

où Ω est l'ensemble des pôles de $G\backslash\{e\}$.

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 (19.619). D'abord, pour toute valeur de $|G|$ (strictement plus grande que 1),

$$1 \leqslant 2 - \frac{2}{|G|} < 2. \tag{19.621}$$

Ensuite, si g fixe x alors g^{-1} fixe également x, de sorte que $n_i = |\operatorname{Fix}(x_i)| \geqslant 2$ pour tout i. Donc tous les termes dans la somme à droite de (19.619) sont dans $[\frac{1}{2}, 1[$. Nous avons donc au minimum deux termes, et au maximum trois. Autrement dit : $r = 2$ ou $r = 3$.

Si $r = 2$ Le plus simple est de repartir de (19.620). 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 $\mathrm{SO}(2)$ et donc cyclique par le lemme 19.141.

Les possibilités pour $r = 3$ Nous devons voir les solutions entières $(n_1, n_2, n_3, |G|)$ de

$$2 - \frac{2}{|G|} = 3 - \frac{1}{n_1} - \frac{1}{n_2} - \frac{1}{n_3} < 2. \tag{19.622}$$

Il faut en particulier que

$$\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} > 1, \tag{19.623}$$

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. Vu que $n_i = |\operatorname{Fix}(x_i)| \geqslant 2$, 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.

Si $n_2 = 2$ Nous sommes dans le cas $r = 3$, $n_1 = 2$, $n_2 = 2$. Nous avons

$$\frac{1}{n_3} = \frac{2}{|G|}, \tag{19.624}$$

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 $\operatorname{Fix}(x)$ est un groupe à $|G|/2$ éléments. Il est donc normal par le lemme 3.29. Si $g \in G$ est tel que $g(x) = y$ alors nous avons $\operatorname{Fix}(y) = g\operatorname{Fix}(x)g^{-1}$, mais comme $\operatorname{Fix}(x)$ est normal nous avons $\operatorname{Fix}(x) = \operatorname{Fix}(y)$. Donc tous les éléments de $\operatorname{Fix}(x)$ fixent x et y. Le groupe $\operatorname{Fix}(x)$ est donc un sous-groupe de $\mathrm{SO}(2)$ est 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 $\mathrm{Fix}(x)$ 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 vu que $s(x) = y$, l'injectivité de s donne $s(y) = x$.

Soit a, un générateur de $\mathrm{Fix}(x)$. Nous allons montrer que $G = \mathrm{gr}(s, sa)$. Nous avons déjà

$$(sa)(x) = s(x) = y \tag{19.625a}$$
$$(sa)(y) = s(y) = x. \tag{19.625b}$$

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 deux 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 $\mathrm{Fix}(x)$. Mais $a = s \cdot sa$, donc $a^k = (ssa)^k$. Nous en déduisons que $\mathrm{gr}(s, sa)$ contient au moins $\mathrm{Fix}(x)$.

D'autre part si h et h' sont des éléments distincts dans $\mathrm{Fix}(x)$, alors sh et sh' sont des éléments distincts de $\mathrm{gr}(s, sa)$ qui ne sont pas dans $\mathrm{Fix}(x)$. Autrement dit, la partie

$$A = \{sh \text{ tel que } h \in Fix(x)\} \tag{19.626}$$

est une partie de même cardinal que $\mathrm{Fix}(x)$ tout en n'ayant aucune intersection avec $\mathrm{Fix}(x)$ (note : l'identité n'est pas dans A). Mais $|\mathrm{Fix}(x)| = |G|/2$, donc $A \cup \mathrm{Fix}(x) = G$. Et justement $A \cup G \subset \mathrm{gr}(s, sa)$. Nous en déduisons que $\mathrm{gr}(s, sa) = G$.

Le théorème 19.105 nous assure que le groupe G est alors le groupe diédral parce que les éléments s et sa vérifient les relations données en 19.99.

Si $r = 3$, les autres cas possibles Nous repartons de (19.619) en posant $r = 3$. Nous obtenons ceci :

$$1 + \frac{2}{|G|} = \frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}. \tag{19.627}$$

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 vallent 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

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{4} = 1 < 1 + \frac{2}{|G|}. \tag{19.628}$$

Donc $n_3 = 3$ est obligatoire. Nous avons alors l'inégalité suivante qui restreint n_3 :

$$\frac{1}{n_3} = \frac{1}{6} + \frac{3}{|G|} > \frac{1}{6}. \tag{19.629}$$

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, $|G| = 60$.

Le cas $(2, 3, 3)$ Nous utilisons les relations $n_i|\Omega_i| = |G|$ pour savoir la taille des orites. Nous avons :

(1) $2|\Omega_1| = 12$, donc $|\Omega_1| = 6$, $3|\Omega_2| = 12$, donc $|\Omega_2| = 4$, $3|\Omega_3| = 12$, 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 $g\big(g^{-1}(x)\big) = x$; donc Ω_2 est bien dans l'image de Ω_2 par G. Nous avons donc un morphisme $s\colon G \to S_{\Omega_2}$ que nous allons immédiatement voir comme

$$s\colon G \to S_4 \tag{19.630}$$

où S_4 est le groupe des permutations de $\{1,2,3,4\}$.

Voyons que s est injective. Si $s(g)=s(h)$, alors $s(gh^{-1})=\mathrm{Id}$. Autrement dit, l'élément sh^{-1} de G est l'identité sur Ω_2 qui contient 4 éléments. Fixant 4 points (au moins), l'élément sh^{-1} est l'identité. Par conséquent

$$s\colon G\to s(G)\subset S_4 \tag{19.631}$$

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 5.30(3).

Le cas $(2,3,4)$ Nous avons $|G=24|$ et les orbites ont pour taille :

— $2|\Omega_1|=24$, donc $|\Omega_1|=12$,

— $3|\Omega_2|=24$, donc $|\Omega_2|=8$,

— $4|\Omega_3|=24$, donc $|\Omega_3|=6$.

Ω_2 vient par paires Soit $x\in\Omega$ tel que $|\operatorname{Fix}(x)|=3$. Alors $x\in\Omega_2$ parce que x est forcément dans un des Ω_i et tout élément x_i de Ω_i vérifie $|\operatorname{Fix}(x_i)|=n_i$. Mais comme les éléments de $\mathrm{SO}(3)$ 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

$$\Omega_2=\{\pm a_1,\pm a_2,\pm a_3,\pm a_4\}. \tag{19.632}$$

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 $g\big(\{a_i,-a_i\}\big)=\{a_k,-a_k\}$. Cette association $i\mapsto k$ est bijective (sinon g ne serait pas bijective), et fournit donc un morphisme de groupes

$$s\colon G\to S_4. \tag{19.633}$$

s est injective Nous prouvons à présent que $s(g)=\mathrm{Id}$ si et seulement si $g=e$. Dans un sens c'est évident : $g(e)=\mathrm{Id}$. 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 Ω_2 ; donc $g^2=e$. Nommons $\pm b$ les points fixes de g. Si $b\in\Omega_2$ alors $|\operatorname{Fix}(b)|=3$, c'est-à-dire que les éléments de G qui fixent b sont dans un groupe d'ordre 3, et le corollaire 2.32 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 Ω_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 $h\in G$. Vu que Ω_2 est une orbite, $h(a_i)\in\Omega_2$ et nous notons $h(a_i)=\epsilon a_k$ avec $\epsilon=\pm1$ et éventiellement $k=i$ ou éventiellement pas. Nous avons :

$$(h^{-1}gh)(a_i)=\epsilon(h^{-1}g)(a_k)=-\epsilon h^{-1}(a_k)=-\epsilon^2a_i=-a_i. \tag{19.634}$$

Donc g et $h^{-1}gh$ ont même restriction à Ω_2. En particulier $h^{-1}ghg^{-1}$ est l'identité sur Ω_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 $h\to h^{-1}$). Cela signifie que $\{b,-b\}$ est une orbite de G. Maizon'a pas d'orbites de cardinal deux ; contradition. Nous en déduisons que e est l'unique élément de $\ker(s)$.

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 [52], nous sautons un niveau d'indentation.

Au niveau du cardinal de G,

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} = 1 + \frac{2}{|G|}, \tag{19.635}$$

donc $|G| = 60$. Et pour les orbites, $|\Omega_1| = 30$, $|\Omega_2| = 20$, $|\Omega_3| = 12$.

La proposition 5.40 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 ?

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 (corollaire 2.32), tous les éléments de G ont un ordre 2, 3 ou 5. Nous sommes dans un cas très particuliers 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.

$g \in G(n_i)$ 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 $x \in Fix(g)$ signifie $g(x) = x$ et donc aussi $g \in \text{Fix}(x)$. Donc l'ordre de g divise $|Fix(x)|$, 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 $x \in \Omega_i$.

$|\text{Fix}(x)| = n_i$ 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 $x \in \Omega_i$.

a et $-a$ dans la même orbite Nous avons évidemment $\text{Fix}(a) = \text{Fix}(-a)$ du fait 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 aussi $-a \in \Omega_i$.

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,

$$\text{Fix}(a) \cap \text{Fix}(b) = \begin{cases} \text{Fix}(a) & \text{si } a = \pm b \\ \{e\} & \text{sinon.} \end{cases} \tag{19.636}$$

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 $x_i \in \Omega_i$.

Récapitulatif En reprenant ce que nous venons de dire avec $i = 1, 2, 3$ nous trouvons :

(1) $n_1 = 2$, avec $|\Omega_1| = 30$ et 15 groupes du type $\text{Fix}(x_1)$ avec x_1 parcourant Ω_1.

(2) $n_2 = 3$, avec $|\Omega_2| = 20$ et 10 groupes du type $\text{Fix}(x_2)$ avec x_2 parcourant Ω_2.

(3) $n_3 = 5$, avec $|\Omega_3| = 12$ et 6 groupes du type $\text{Fix}(x_3)$ avec x_3 parcourant Ω_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 pour la seconde ligne donn 10 groupes de taille $|\text{Fix}(x_2)| = n_2 = 3$. Mais

52. Mais pas autant que le théorème 33.28, cependant.

tout ces groupes ont pour intersection exactement $\{e\}$. Donc le comptage des éléments se fait comme suit :

$$3 \times 30 - 15 - 10 - 6 + 1. \tag{19.637}$$

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.

Un ensemble à calculer deux fois Soient les ensembles A_2, A_3 et A_5 définis par

$$A_i = \{(g, a) \in G(n_i) \times \Omega_i \text{ tel que } g(a) = a\} \tag{19.638}$$

où $G(n_i)$ est la partie de G des éléments d'ordre n_i.

Nous avons

$$|A_i| = \sum_{g \in G(n_i)} |\operatorname{Fix}(g) \cap \Omega_i|. \tag{19.639}$$

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 $\operatorname{Fix}(g)$ sont dans Ω_i. Donc $\operatorname{Fix}(g) \cap \Omega_i = \operatorname{Fix}(g)$. Nous avons alors

$$|A_i| = \sum_{g \in G(n_i)} |\operatorname{Fix}(g)| = 2|G(n_i)| \tag{19.640}$$

parce que $|\operatorname{Fix}(g)| = 2$ pour tout g.

En compatant $|A_i|$ dans l'autre sens, nous avons

$$|A_i| = \sum_{x \in \Omega_i} |\operatorname{Fix}(x) \cap G(n_i)| \tag{19.641}$$

Vu que $x \in \Omega_i$, les éléments de $\operatorname{Fix}(x)$ sont d'ordre n_i [53] (sauf e), et comme $G(n_i)$ est justement l'ensemble des éléments d'ordre n_i dans G nous avons $\operatorname{Fix}(x) \cap G(n_i) : \operatorname{Fix}(x) \backslash \{e\}$. Cela pour dire que

$$\begin{aligned}
|A_i| &= \sum_{x \in \Omega_i} \Big(|\operatorname{Fix}(x)| - 1 \Big) && && \text{(19.642a)}\\
&= \sum_{x \in \Omega_i} |\operatorname{Fix}(x)| - \sum_{x \in \Omega_i} 1 && && \text{(19.642b)}\\
&= \sum_{x \in \Omega_i} n_i - |\Omega_i| && |\operatorname{Fix}(x)| = n_i \text{ pcq } x \in \Omega_i && \text{(19.642c)}\\
&= |\Omega_i| n_i - |\Omega_i| = |G| - |\Omega_i|. && && \text{(19.642d)}
\end{aligned}$$

En égalisant cela à la valeur $2|G(n_i)|$ déjà trouvée, nous déduisons les valeurs des $|G(n_i)|$:

$$|G(n_i)| = \frac{|G| - |\Omega_i|}{2}. \tag{19.643}$$

Nous avons alors

(1) $|G(2)| = 15$
(2) $|G(3)| = 20$
(3) $|G(5)| = 24$

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 :

(1) Un 2-Sylow est un sous-groupe d'ordre 4.

53. Encore et toujoure parce que les éléments de $\operatorname{Fix}(x)$ ont un ordre qui divise $|\operatorname{Fix}(x)| = n_i$ et que n_i est premier, et que nous avons exclu l'identité.

(2) Un 3-Sylow est un sous-groupe d'ordre 3.

(3) Un 5-Sylow est un sous-groupe d'ordre 5.

Entre autres :

(1) Les 10 sous-groupes $\mathrm{Fix}(x_2)$ avec $x_2 \in \Omega_2$ sont des 3-Sylow de G.

(2) Les 6 sous-groupes $\mathrm{Fix}(x_3)$ avec $x_3 \in \Omega_3$ sont des 5-Sylow de G.

(3) Les 15 sous-groupes $\mathrm{Fix}(x_1)$ avec $x_1 \in \Omega_1$ sont d'ordre 2 et ne sont donc pas des 2-Sylow de G.

Tous les 3-Sylow et les 5-Sylow Nous avons déjà trouvé 10 3-Sylow et 6 5-Sylow. Nous motrons 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 :

$$1, 5, 3, 15, 2, 10, 6, 30, 4, 20, 12, 60. \tag{19.644}$$

Le seul qui vérifie toutes les conditions est 10. Donc G possède seulement 10 3-Sylow et ils sont tous de la forme $\mathrm{Fix}(x_2)$ 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.

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 2.31(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.

Diviseurs de $|H|$ Soit un sous-groupe normal H de G. Vu que c'est un sous-groupe sont ordre divise celui de G (encore et toujours le théorème de Lagrange 2.31), et donc les facteurs premiers de $|H|$ ne peuvent être que 2, 3 et 5.

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-Syow sont conjugués et H est normal.

Soit E l'ensemble des sous-groupes de H. Vu qu'il est normal, H agit sur E par conjugaison, et les 3-Sylow forment une orbite. Si α est un 3-Sylow, la formule des classes (proposition 2.49(2)) nous donne

$$|H| = |\mathrm{Fix}(\alpha)||\mathcal{O}_\alpha|. \tag{19.645}$$

Mais l'orbite $\mathcal{O}_\alpha$ de α 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 $\mathrm{Fix}(x_2)$ pour $x_2 \in \Omega_2$. Ce sont 10 groupes de $|\mathrm{Fix}(x_2)| = 3$ éléments. En décomptant e qui est dans l'intersection, cela fait

$$10 \times |\mathrm{Fix}(x_2)| - 10 + 1 = 21 \tag{19.646}$$

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.

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.

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 $\mathrm{Fix}(x_1)$ sont dans l'unique 2-Sylow H qui est soit-disant d'ordre 4. IL y a là une belle impossibilité.

Donc le cas $|H| = 4$ est hors-concours.

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.

Vu que $g \in G(2)$, nous avons que les éléments $a \in \mathrm{Fix}(g)$ sont forcément dans Ω_1 parce que les points dont les fixateurs sont formés d'éléments d'ordre 2 sont dans Ω_1. Soit $h \in G$. Nous avons $g = hgh^{-1}$ et donc aussi

$$\big(hgh^{-1}\big)\big(h(a)\big) = hg(a) = h(a), \tag{19.647}$$

donc $h(a)$ et $-h(a)$ sont des points fixes de hgh^{-1}. Ce sont donc également de points fixes de g. Nous en déduisons que g a pour points fixes les points a, $-a$, $h(a)$ et $-h(a)$. Vu que g n'est pas e, ces quatre points ne peuvent pas être distincts. Vu que $h(a)$ 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.

Si $|H| = 30$ À part $|H| = 60$, le dernier cas à traiter est $|H| = 30$. Nous rappellons obligeament que

(1) $|G(2)| = 15$

(2) $|G(3)| = 20$

(3) $|G(5)| = 24$.

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. Vu que pour 3 et 5, les Sylow de H et de G sont les mêmes et bien idéntifiés, nous allons nous baser dessus. Le sous-groupe H contient tous les 3 et 5-Sylow, donc le comptage des éléments est :

$$10 \times |\mathrm{Fix}(x_2)| + 6 \times |\mathrm{Fix}(x_3)| - 15 = 45. \tag{19.648}$$

Nous aurions aussi pu ajouter $+4 - 1$ pour compter au moins un 2-Sylow.

Donc dès que H compte 30 éléments, il en compte 45 et donc 60 parce qu'il n'y a pas de diviseurs de 60 entre 45 et 60.

□

Problèmes et choses à faire

La démonstration des groupes finis de SO(3) est longue. Je me demande si il n'y a pas moyen de faire plus court. Par exemple [329] utilise le théorème de Cauchy 3.27 que je n'utilise pas. D'autre part, toutes les références me semble 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.

19.12 Systèmes de coordonnées

La trigonométrique nous offre de nouveaux systèmes de coordonnées qui peuvent se révéler pratique de certains cas : les coordonnées polaires sur $\mathbb{R}^2$ ainsi que les coordonnées cylindriques et sphériques sur $\mathbb{R}^3$.

19.12.1 Coordonnées polaires

19.12.1.1 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 montré sur la figure 19.11.

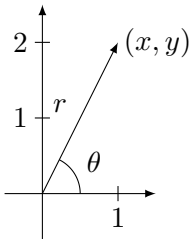


FIGURE 19.11 – Un point en coordonnées polaires est donné par sa distance à l'origine et par l'angle qu'il faut avec l'horizontale.

Le même point M peut être décrit indifféremment avec les coordonnées (x, y) ou bien avec (r, θ).

Remarque 19.189.
L'angle θ d'un point n'étant a priori défini qu'à un multiple de 2π près, nous convenons de toujours choisir un angle $0 \leqslant \theta < 2\pi$. Par ailleurs l'angle θ n'est pas défini si $(x, y) = (0, 0)$.

La coordonnée r est toujours positive.

Nous avons dans l'idée de définir r et θ par les formules

$$\begin{cases} x = r\cos(\theta) & (19.649a) \\ y = r\sin(\theta). & (19.649b) \end{cases}$$

19.12.1.2 Coordonnées polaires : le théorème

Théorème 19.190 (Coordonnées polaires[1])**.**
Soit l'application

$$\begin{aligned} T\colon [0,\infty[\times [0, 2\pi[&\to \mathbb{R}^2 \\ (r,\theta) &\mapsto \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \end{aligned} \tag{19.650}$$

(1) L'application T est surjective.

(2) L'application

$$T\colon]0,\infty[\times [0, 2\pi[\to \mathbb{R}^2\backslash\{(0,0)\} \tag{19.651}$$

est bijective.

(3) En considérant la demi-droite $D = \{(x, 0)\}_{x\geqslant 0}$, l'application

$$T\colon]0,\infty[\times]0, 2\pi[\to \mathbb{R}^2\backslash D \tag{19.652}$$

est un C^∞-difféomorphisme[54].

54. L'application est de classe C^∞ et son inverse est également de classe C^∞. 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.

Démonstration. Une bonne partie de ce théorème est une conséquence de 19.58. 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 19.58, il existe $\theta \in [0, 2\pi]$ tel que

$$\frac{a}{\|a\|} = \big(\cos(\theta), \sin(\theta)\big). \tag{19.653}$$

Alors $a = \|a\|\big(\cos(\theta), \sin(\theta)\big) = T(\|a\|, \theta)$. Voila. L'application T est surjective.

(2) Pour (2). En ce qui concerne la surjectifivé,

$$T\big(0, [2, 2\pi[\big) = \{(0, 0)\}. \tag{19.654}$$

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 $T(r_1, \theta_1) = T(r_2, \theta_2)$. Vu que $\|T(t, \theta)\| = r$, nous avons tout de suite $r_1 = r_2$. Nous restons donc avec l'égalité

$$\begin{pmatrix} \cos(\theta_1) \\ \sin(\theta_1) \end{pmatrix} = \begin{pmatrix} \cos(\theta_2) \\ \sin(\theta_2) \end{pmatrix}. \tag{19.655}$$

La proposition 19.58 dit alors que $\theta_1 = \theta_2$.

(3) Pour (3). L'application T est injective en tant que restriction d'une application injective. Pour le surjectif, soit $a \in \mathbb{R}^2 \backslash D$. Vu que $a \notin D$, nous avons $\|a\| \neq 0$ et il est légitime de dire, comme plus haut, qu'il existe $\theta \in [0, 2\pi[$ tel que

$$\frac{a}{\|a\|} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}. \tag{19.656}$$

Ce θ 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 (19.170) ; nous avons déjà démontré la continuité dans le lemme 19.62, et monter dans les dérivées nous semble un peu compliqué. Au lieu de cela, nous allons faire 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 18.48.

Les applications $(r, \theta) \mapsto r$, $(r, \theta) \mapsto \sin(\theta)$ et $(r, \theta) \mapsto \cos(\theta)$ sont de classe C^∞ grâce au lemme 13.189. Le lemme 13.212 sur la différentiabilité du produit montre alors que les fonctions T_1 et T_2 données par

$$T_1(r, \theta) = r\cos(\theta) \tag{19.657a}$$
$$T_2(r, \theta) = r\sin(\theta) \tag{19.657b}$$

sont différentiables [55]. Mieux, la proposition 13.215 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^∞. Cela montre que les coordonnées polaires sont de classe C^∞, et il faut encore parler de l'inverse.

En ce qui concerne la différentielle,

$$dT_{(r,\theta)}(u, v) = \begin{pmatrix} u\cos(\theta) - rv\sin(\theta) \\ u\sin(\theta) + rv\cos(\theta) \end{pmatrix}. \tag{19.658}$$

55. 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 13.215 ni rien des produits tensoriels.

Donc la matrice de la différentielle est

$$dT_{(r,\theta)} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta), \end{pmatrix} \tag{19.659}$$

dont le déterminant est r (lemme 19.18 utilisé). Donc la différentielle en (r,θ) 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 18.48 fonctionne et T^{-1} est C^p dès que T est C^p.

Vu que 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^∞.

□

Définition 19.191.
Ce que nous appelons ***les coordonnées polaires*** *est l'application*

$$\begin{aligned} T\colon [0,\infty[\times [0,2\pi[&\to \mathbb{R}^2 \\ (r,\theta) &\mapsto \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \end{aligned} \tag{19.660}$$

du théorème 19.190(3). Selon les circonstances, nous considérons l'une ou l'autre des restrictions pour avoir une bijection ou un difféomorphisme.

Exemple 19.192
Soit à calculer

$$\lim_{(x,y)\to(0,0)} \frac{x^2+y^2}{x-y}. \tag{19.661}$$

Nous introduisons la fonction

$$\begin{aligned} f\colon \mathbb{R}^2\backslash\{x=y\} &\to \mathbb{R} \\ (x,y) &\mapsto \frac{x^2+y^2}{x-y}. \end{aligned} \tag{19.662}$$

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

$$\begin{aligned} f\colon \big(]0,\infty[\times [0,2\pi[\big)\backslash\big\{\mathbb{R}\times\{\frac{\pi}{4}\}\cup\mathbb{R}\times\{\frac{5\pi}{4}\}\big\} &\to \mathbb{R} \\ (r,\theta) &\mapsto \frac{r^2}{r\big(\cos(\theta)-\sin(\theta)\big)}. \end{aligned} \tag{19.663}$$

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 $r\to 0$, 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

$$g = f \circ T \tag{19.664}$$

où T est l'application de coordonnées polaires dont les principales propriétés sont données dans le théorème 19.190(2).

Soit un voisinage $B\big((0,0),R\big)$ de $(0,0)$ dans $\mathbb{R}^2$. Le but est de montrer que les valeurs $f(B)$ se regroupent autour d'une valeur ℓ lorsque $R\to 0$. Soyons plus précis et nommons ℓ le candidat limite. Soit $\epsilon>0$; nous devons trouver $R>0$ tel que $f\Big(B\big((0,0),R\big)\Big)\subset B(\ell,\epsilon)$.

Pour $R > 0$, nous avons

$$B\big((0,0),R\big) = T\big([0,R[\times [0,2\pi[\big), \tag{19.665}$$

donc

$$f(B) = g\big([0,R[\times [2,2\pi[\big). \tag{19.666}$$

Soit $r < R$. Nous avons

$$\lim_{\theta\to\pi/4} g(r,\theta) = \infty. \tag{19.667}$$

Donc $f(B)$ contient des valeurs arbitrairement grandes, quelle que soit la valeur de R. Il n'y a donc pas de limite possibles.

Si vous voulez un argument un peu plus imagé, en voici un [56] basé sur une combinaison entre la méthode des coordonnées polaires et la méthode des chemins.

Certes *pour chaque* θ nous avons $\lim_{r\to 0} g(r,\theta) = 0$, mais il ne faut pas en déduire trop vite que la limite $\lim_{(x,y)\to(0,0)} g(x,y)$ vaut zéro parce que prendre la limite $r \to 0$ avec θ fixé revient à prendre la limite le long de la droite d'angle θ.

Il n'est pas possible de majorer $g(r,\theta)$ par une fonction ne dépendant pas de θ parce que cette fonction tend vers l'infini lorsque $\theta \to \pi/4$. 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.

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

$$\cos\theta(t) - \sin\theta(t) = t^2. \tag{19.668}$$

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 θ qui résout l'équation (19.668).

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 18.50. Nous verrons dans l'exemple 19.207 comment s'en sortir sans théorème de la fonction implicite, au prix de plus de calculs. △

Les coordonnées polaires sont données par le difféomorphisme

$$\begin{aligned} g\colon]0,\infty[\times]0,2\pi[&\to \mathbb{R}^2\backslash D \\ (r,\theta) &\mapsto \big(r\cos(\theta), r\sin(\theta)\big) \end{aligned} \tag{19.669}$$

où D est la demi-droite $y = 0$, $x \geqslant 0$. 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

$$Jg = \det\begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \det\begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} = r. \tag{19.670}$$

La fonction qui donne les coordonnées polaires est

$$\begin{aligned} \varphi\colon \mathbb{R}^+ \times]0,2\pi[&\to \mathbb{R}^2 \\ (r,\theta) &\mapsto \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \end{aligned} \tag{19.671}$$

Son Jacobien vaut

$$J_\varphi(r,\theta) = \det\begin{pmatrix} \frac{\partial x(r,\theta)}{\partial r} & \frac{\partial x(r,\theta)}{\partial \theta} \\ \frac{\partial y(r,\theta)}{\partial r} & \frac{\partial y(r,\theta)}{\partial \theta} \end{pmatrix} = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{vmatrix} = r. \tag{19.672}$$

56. Qui satisfera tous vos professeurs, pourvu que vous ayez compri que ce qui se cache est une histoire de valeurs de f prises sur un voisinage de $(0,0)$.

Proposition 19.193.
Soit la fonction

$$\begin{aligned} T\colon]O,+\infty[\times\mathbb{R} &\to \mathbb{R}^2\backslash\{(0,0)\}\\ (r,\theta) &\mapsto \big(r\cos(\theta), r\sin(\theta)\big). \end{aligned} \tag{19.673}$$

(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

$$T^{-1}(x,y)=\big(\sqrt{x^2+y^2},\operatorname{arctg}(y/x)\big). \tag{19.674}$$

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 $\theta=\operatorname{arctg}(y/x)$ 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 θ.

Exemple 19.194
L'équation du cercle de rayon a et centre $(0,0)$ en coordonnées polaires est $r=a$. △

Exemple 19.195
Une équation possible pour la demi-droite $x=y$, $x>0$, est $\theta=\pi/4$. △

19.12.1.3 Transformation inverse : théorie

Voyons la question inverse : comment retrouver r et θ si on connait x et y ? Tout d'abord,

$$r=\sqrt{x^2+y^2} \tag{19.675}$$

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

$$\theta=\begin{cases}\pi/2 & \text{si } y>0\\ 3\pi/2 & \text{si } y<0\end{cases} \tag{19.676}$$

Notez que si $y<0$, conformément à notre convention $\theta\geqslant 0$, nous avons noté $\frac{3\pi}{2}$ et non $-\frac{\pi}{2}$.

Supposons maintenant le cas général avec $x\neq 0$. Les équations (19.649) montrent que

$$\tan(\theta)=\frac{y}{x}. \tag{19.677}$$

Nous avons donc

$$\theta=\tan^{-1}\left(\frac{y}{x}\right). \tag{19.678}$$

La fonction inverse de la fonction tangente est celle définie plus haut.

19.12.1.4 Transformation inverse : pratique

Le code suivant utilise Sage.

```
#! /usr/bin/sage -python
# -*- coding: utf8 -*-

from sage.all import *

def PointToPolaire(x,y):
```

```
  x=SR(x)
  y=SR(y)
  r = sqrt(x**2+y**2)
  if x == 0:
    if y > 0:
      alpha = pi/2
    if y < 0:
      alpha = 3*pi/2
    if y == 0 :
      raise ValueError,"Pas d'angle pour le point (0,0) !!"
  else :
    alpha = atan(y/x)
  if (x < 0) and (y == 0) :
    alpha = pi
  if (x < 0) and (y > 0) :
    alpha = alpha + pi
  if (x < 0) and (y < 0 ) :
    alpha = alpha + pi
  alpha=alpha.simplify_trig()
  return (r,alpha)

print PointToPolaire(1,1)
print PointToPolaire(-2,1)
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.

19.12.1.5 Coordonnées polaires : dérivées partielles

Le changement de coordonnées pour les polaires est la fonction

$$f\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}. \tag{19.679}$$

Considérons une fonction g sur $\mathbb{R}^2$, et définissons la fonction $\tilde{g}$ par

$$\tilde{g}(r,\theta) = g(r\cos\theta, r\sin\theta). \tag{19.680}$$

La formule (13.479) permet de trouver les dérivées partielles de g par rapport à r et θ en termes de celles par rapport à x et y de g.

Pour faire le lien avec les notations du point précédent, nous avons

$$\begin{aligned} f_1(r,\theta) &= r\cos(\theta) \\ f_2(r,\theta) &= r\sin(\theta) \\ (x_1,x_2) &\to (r,\theta) \\ (y_1,y_2) &\to (x,y). \end{aligned} \tag{19.681}$$

Nous avons donc

$$\begin{aligned}\frac{\partial\tilde{g}}{\partial r}(r,\theta) &= \sum_{i=1}^{2}\frac{\partial g}{\partial x_i}\big(f(r,\theta)\big)\frac{\partial f_i}{\partial r}(r,\theta)\\ &= \frac{\partial g}{\partial x}(r\cos\theta, r\sin\theta)\frac{\partial\big(r\cos\theta\big)}{\partial r}(r,\theta)\\ &\quad+\frac{\partial g}{\partial y}(r\cos\theta, r\sin\theta)\frac{\partial\big(r\sin\theta\big)}{\partial r}(r,\theta)\\ &= \cos\theta\frac{\partial g}{\partial x}(r\cos\theta, r\sin\theta)+\sin\theta\frac{\partial g}{\partial y}(r\cos\theta, r\sin\theta).\end{aligned}\tag{19.682}$$

Prenons par exemple $g(x,y)=\frac{1}{x^2+y^2}$. Étant donné que

$$\frac{\partial g}{\partial x}=\frac{-2x}{(x^2+y^2)^2},\tag{19.683}$$

nous avons

$$\frac{\partial g}{\partial x}(r\cos\theta, r\sin\theta)=\frac{-2\cos\theta}{r^3}.\tag{19.684}$$

En utilisant la formule,

$$\frac{\partial\tilde{g}}{\partial r}(r,\theta)=\cos(\theta)\left(\frac{-2\cos\theta}{r^3}\right)+\sin(\theta)\left(\frac{-2\sin\theta}{r^3}\right)=-\frac{2}{r^3}.\tag{19.685}$$

Nous pouvons vérifier directement que cela est correct. En effet

$$\tilde{g}(r,\theta)=g(r\cos\theta, r\sin\theta)=\frac{1}{r^2},\tag{19.686}$$

dont la dérivée par rapport à r vaut $-2/r^3$.

En ce qui concerne la dérivée par rapport à θ, nous avons

$$\begin{aligned}\frac{\partial\tilde{g}}{\partial\theta} &= \frac{\partial g}{\partial x}(r\cos\theta, r\sin\theta)\frac{\partial\big(r\cos(\theta)\big)}{\partial\theta}+\frac{\partial g}{\partial y}(r\cos\theta, r\sin\theta)\frac{\partial\big(r\sin(\theta)\big)}{\partial\theta}\\ &= \left(\frac{-2\cos\theta}{r^3}\right)(-r\sin\theta)+\left(\frac{-2\sin\theta}{r^3}\right)(r\cos\theta)\\ &= 0.\end{aligned}\tag{19.687}$$

En résumé et avec quelques abus de notation :

$$\begin{aligned}\frac{\partial\tilde{g}}{\partial r} &= \cos(\theta)\frac{\partial g}{\partial x}+\sin(\theta)\frac{\partial g}{\partial y}\\ \frac{\partial\tilde{g}}{\partial\theta} &= -r\sin(\theta)\frac{\partial g}{\partial x}+r\cos(\theta)\frac{\partial g}{\partial y}\end{aligned}\tag{19.688}$$

19.12.2 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

$$\begin{cases}x=r\cos(\theta) & \text{(19.689a)}\\ y=r\sin(\theta) & \text{(19.689b)}\\ z=z. & \text{(19.689c)}\end{cases}$$

Soit T la fonction de $]0,+\infty[\times\mathbb{R}^2$ dans $\mathbb{R}^3\backslash\{(0,0,0)\}$ définie par

$$\begin{array}{rccc}T: &]0,+\infty[\times\mathbb{R}\times\mathbb{R} & \to & \mathbb{R}^3\backslash\{(0,0,0)\}\\ & (r,\theta,z) & \mapsto & (r\cos\theta, r\sin\theta, z),\end{array}\tag{19.690}$$

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 19.196
Il faut savoir ce que décrivent les équations les plus simples en coordonnées cylindriques,

— $r \leqslant a$, 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 x-y.

△

Exemple 19.197
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 19.198
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

$$J(r,\theta,z) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r. \tag{19.691}$$

Nous avons donc $dx\,dy\,dz = r\,dr\,d\theta\,dz$.

$$\begin{cases} x = r\cos\theta & \text{(19.692a)} \\ y = r\sin\theta & \text{(19.692b)} \\ z = z & \text{(19.692c)} \end{cases}$$

avec $r \in]0,\infty[$, $\theta \in [0,2\pi[$ et $z \in \mathbb{R}$. Le jacobien vaut r.

19.12.3 Coordonnées sphériques

Soit T la fonction de $]0,+\infty[\times\mathbb{R}^2$ dans $\mathbb{R}^3\backslash\{(0,0,0)\}$ définie par

$$\begin{array}{rccc} T: &]0,+\infty[\times\mathbb{R}\times\mathbb{R} & \to & \mathbb{R}^3\backslash\{(0,0,0)\} \\ & (\rho,\theta,\phi) & \mapsto & (\rho\cos\theta\sin\phi, \rho\sin\theta\sin\phi, \rho\cos\phi), \end{array} \tag{19.693}$$

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 donnée

$$\begin{array}{rccc} T: & \mathbb{R}^3\backslash\{(0,0,0)\} & \to &]0,+\infty[\times[-\pi,\pi[\times[0,\pi[\\ & (x,y,z) & \mapsto & \left(\sqrt{x^2+y^2+z^2}, \operatorname{arctg}\frac{y}{x}, \arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right)\right). \end{array} \tag{19.694}$$

Soit P un point dans $\mathbb{R}^3$. L'angle ϕ est l'angle entre le demi-axe positif des z et le vecteur $\overrightarrow{OP}$, ρ est la norme de $\overrightarrow{OP}$ et θ est l'argument en coordonnées polaires de la projection de $\overrightarrow{OP}$ sur le plan x-y.

Remarque 19.199.
Dans la littérature, les angles θ et ϕ sont parfois inversés (voire, changent de nom, par exemple φ au lieu de ϕ). Il faut donc être très prudent lorsqu'on veut utiliser dans un cours des formules données dans un autre cours.

Exemple 19.200
Il faut connaitre le sens des équations plus simples,

— $\rho \leqslant a$, 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

$$\begin{cases} x = \rho \sin(\theta) \cos(\varphi) & (19.695a) \\ y = \rho \sin(\theta) \sin(\varphi) & (19.695b) \\ z = \rho \cos(\theta) & (19.695c) \end{cases}$$

avec $0 \leqslant \theta \leqslant \pi$ et $0 \leqslant \varphi < 2\pi$.

Remarque 19.201.
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 φ est noté θ et l'angle θ est noté Φ. Mais vous noterez que sur cette même page, les conventions de noms de ces angles changent plusieurs fois.

Les coordonnées sphériques sont données par

$$\begin{cases} x = r \cos\theta \sin\varphi \\ y = r \sin\theta \sin\varphi \\ z = r \cos\varphi \end{cases} \quad \text{avec} \quad \begin{array}{l} r \in]0, \infty[\\ \theta \in]0, 2\pi[\\ \phi \in]0, \pi[. \end{array} \tag{19.696}$$

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.

$$\begin{cases} x = \rho \cos\theta \sin\phi & (19.697a) \\ y = \rho \sin\theta \sin\phi & (19.697b) \\ z = \rho \cos\phi & (19.697c) \end{cases}$$

avec $\rho \in]0, \infty[$, $\theta \in [0, 2\pi[$ et $\phi \in [0, \pi[$. Le jacobien vaut $-\rho^2 \sin(\phi)$.

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 conséquent. Ce n'est pas le cas de toutes les conventions possibles.

19.12.3.1 Coordonnées sphériques : inverse

Trouvons le changement inverse, c'est-à-dire trouvons ρ, θ et φ en termes de x, y et z. D'abord nous avons

$$\rho = \sqrt{x^2 + y^2 + z^2}. \tag{19.698}$$

Ensuite nous savons que

$$\cos(\theta) = \frac{z}{\rho} \tag{19.699}$$

détermine de façon unique[57] un angle $\theta \in [0, \pi]$. Dès que ρ et θ sont connus, nous pouvons poser $r = \rho \sin\theta$ et alors nous nous trouvons avec les équations

$$\begin{cases} x = r\cos(\varphi) & \text{(19.700a)} \\ y = r\sin(\varphi), & \text{(19.700b)} \end{cases}$$

qui sont similaires à celles déjà étudiées dans le cas des coordonnées polaires.

19.13 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 19.193. De plus, le théorème de la fonction implicite Nous en voyons quelques exemples à présent.

Exemple 19.202(Limite et prolongement par continuité)
La fonction

$$f(x) = \frac{\cos(x) - 1}{x} \tag{19.701}$$

n'est pas définie en $x = 0$.

Nous avons vu dans l'équation (19.56) que $\cos(0) = 1$, donc la limite

$$\lim_{x\to 0} \frac{\cos(x) - 1}{x} \tag{19.702}$$

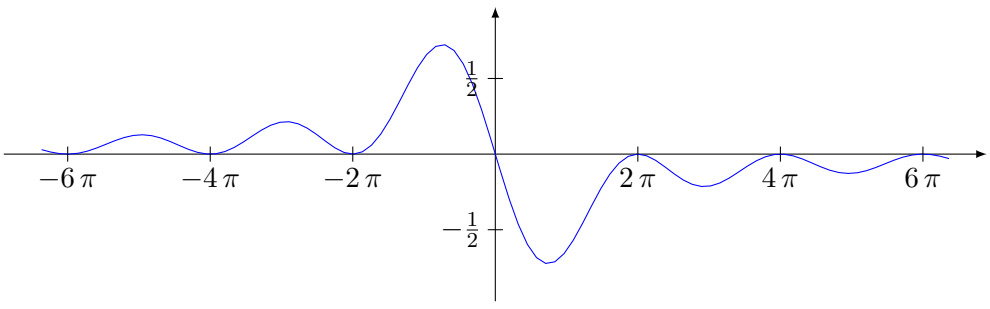
est la limite définissant la dérivée de cosinus en 0 (ici, le x joue le rôle de ϵ). Le lemme 19.17 nous donne la dérivée du cosinus comme étant le sinus. Nous avons donc :

$$\lim_{x\to 0} \frac{\cos(x) - 1}{x} = \sin(0) = 0, \tag{19.703}$$

et nous définissons le prolongement par continuité :

$$\tilde{f}(x) = \begin{cases} \frac{\cos(x)-1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases} \tag{19.704}$$

Encore une fois, le graphe de la fonction $\tilde{f}$ ne présente aucune particularité autour de $x = 0$.



57. Le problème $\rho = 0$ ne se pose pas ; pourquoi ?

△

Exemple 19.203(Un calcul heuristique de limite)
Soit à calculer la limite suivante :

$$\lim_{x\to 0} \frac{e^{-2\cos(x)+2}\sin(x)}{\sqrt{e^{2\cos(x)+2}}-1}. \tag{19.705}$$

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 $y(x)\to 0$. La limite devient

$$\lim_{x\to 0} \frac{\big(-2\cos(x)+3\big)\sin(x)}{\sqrt{-2\cos(x)+2}}. \tag{19.706}$$

Nous allons maintenant remplacer $\cos(x)$ par 1 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

$$\lim_{x\to 0} \frac{\sin(x)}{\sqrt{-2\left(1-\frac{x^2}{2}\right)+2}} = \lim_{x\to 0}\frac{\sin(x)}{x} = 1. \tag{19.707}$$

Tout ceci n'est évidemment pas très rigoureux, mais en principe vous avez tous les éléments en main pour justifier les étapes. △

19.13.1 Méthode des coordonnées polaires

La proposition suivante exprime la définition de la limite en d'autres termes, et va être pratique dans le calcul de certaines limites.

Proposition 19.204.
Soit $f\colon D\subset\mathbb{R}^m\to\mathbb{R}^n$, a un point d'accumulation de D et $\ell\in\mathbb{R}^n$. Nous définissons

$$E_r = \{f(x) \text{ tel que } x\in B(a,r)\cap D\}, \tag{19.708}$$

et

$$s_r = \sup\{\|v-\ell\| \text{ tel que } v\in E_r\}. \tag{19.709}$$

Alors nous avons $\lim_{x\to a} f(x)=\ell$ si et seulement si $\lim_{r\to 0} s_r = 0$.

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 ℓ et donne la distance. En d'autres termes, s_r est la distance maximale entre $f(x)$ et ℓ lorsque x est à une distance au maximum r de a.

Lorsque nous avons affaire à une fonction $f\colon\mathbb{R}^2\to\mathbb{R}$, cette proposition nous permet de calculer facilement les limites en passant aux coordonnées polaires.

Exemple 19.205
Reprenons la fonction de l'exemple 13.154 :

$$f(x,y) = \frac{xy}{x^2+y^2}. \tag{19.710}$$

Son domaine est $\mathbb{R}^2\backslash\{(0,0)\}$. Nous voulons calculer $\lim_{(x,y)\to(0,0)} f(x,y)$. Écrivons la définition de E_r :

$$E_r = \{f(x,y) \text{ tel que } (x,y) \in B\big((0,0),r\big)\}. \tag{19.711}$$

Les points de la boule sont, en coordonnées polaires, les points de la forme (ρ,θ) 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

$$f(\rho,\theta) = \frac{\rho^2\cos(\theta)\sin(\theta)}{\rho^2} = \cos(\theta)\sin(\theta). \tag{19.712}$$

Cela signifie que

$$E_r = \{\cos(\theta)\sin(\theta) \text{ tel que } \theta \in [0,2\pi[\}. \tag{19.713}$$

Prenons ℓ quelconque. Le nombre s_r est le supremum des

$$\|\ell - \cos(\theta)\sin(\theta)\| \tag{19.714}$$

lorsque θ parcours $[0,2\pi]$. Nous ne sommes pas obligés 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 $\lim_{r\to 0} s_r = 0$, et la fonction donnée n'a pas de limite en $(0,0)$. △

Nous pouvons retenir cette règle pour calculer les limites lorsque $(x,y)\to(0,0)$ de fonctions $f\colon\mathbb{R}^2\to\mathbb{R}$:

(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 ρ et θ. Si la limite $\lim_{r\to 0} g(r,\theta)$ n'existe pas ou dépend de θ, alors la fonction n'a pas de limite. Si on peut majorer g par une fonction ne dépendant pas de θ, et que cette fonction a une limite lorsque $r\to 0$, 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 θ. 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 θ.

Exemple 19.206
Soit à calculer la limite

$$\lim_{(x,y)\to(0,0)} xy\left(\frac{x^2-y^2}{x^2+y^2}\right). \tag{19.715}$$

Le passage aux coordonnées polaires donne

$$f(r,\theta) = r^2\sin\theta\cos\theta(\cos^2\theta-\sin^2\theta). \tag{19.716}$$

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

$$\|f(r,\theta)\| \leqslant r^2. \tag{19.717}$$

Maintenant la règle de l'étau montre que $\lim_{(x,y)\to(0,0)} f(x,y)$ est zéro.

La situation vraiment gênante serait celle avec une fonction de θ qui risque de s'annuler dans un dénominateur. △

L'exemple 19.192 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.

Exemple 19.207
Considérons fonction

$$f(x,y) = \frac{x^2+y^2}{x-y}. \tag{19.718}$$

Une mauvaise idée pour prouver que la limite n'existe pas pour $(x,y)\to(0,0)$ 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

$$f(t,kt)=\frac{t(1+k^2)}{1-k}. \tag{19.719}$$

La limite $t\to 0$ 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

$$\frac{1+k(t)^2}{1-k(t)}=\frac{1}{t}. \tag{19.720}$$

Avec une telle fonction k, la fonction $t\mapsto f(t,tk(t))$ serait la constante 1. L'équation à résoudre pour k est

$$tk^2+k+(t-1)=0, \tag{19.721}$$

et les solutions sont

$$k(t)=\frac{-1\pm\sqrt{1-4t(t-1)}}{2t}. \tag{19.722}$$

Nous proposons donc les chemins

$$\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}t\\ \frac{-1\pm\sqrt{1-4t(t-1)}}{2}\end{pmatrix} \tag{19.723}$$

Nous devons vérifier deux points. D'abord que ce chemin est bien défini, et ensuite que $tk(t)$ tend bien vers zéro lorsque $t\to 0$ (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,

$$\lim_{t\to 0}tk(t)=\frac{-1\pm 1}{2}. \tag{19.724}$$

En choisissant le signe $+$, nous trouvons un chemin qui nous convient.

Ce que nous avons prouvé est que

$$f\left(t,\frac{-1+\sqrt{1-4t(t-1)}}{2}\right)=1 \tag{19.725}$$

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 $\lim_{(x,y)\to(0,0)}f(x,y)$ n'existe donc pas. △

Exemple 19.208

Considérons la fonction (figure 19.12)

$$f(x,y)=\begin{cases}\sqrt{x^2+y^2}\sin\frac{1}{x^2+y^2} & \text{si } (x,y)\neq(0,0)\\ 0 & \text{si } (x,y)=(0,0),\end{cases} \tag{19.726}$$

et cherchons la limite $(x,y)\to(0,0)$. Le passage en coordonnées polaires[58] donne

$$f(\rho,\theta)=\rho\sin\frac{1}{\rho}. \tag{19.727}$$

58. Proposition 19.193.

Pour calculer la limite de cela lorsque $\rho \to 0$, nous remarquons que

$$0 \leqslant |\rho \sin \frac{1}{\rho}| \leqslant \rho \tag{19.728}$$

parce que $\sin(\frac{1}{\rho}) \leqslant 1$ quel que soit ρ. Or évidemment $\lim_{\rho\to 0} \rho = 0$, donc la limite de la fonction (19.727) est zéro et ne dépend pas de θ. Nous en concluons que $\lim_{(x,y)\to(0,0)} f(x,y) = 0$. △

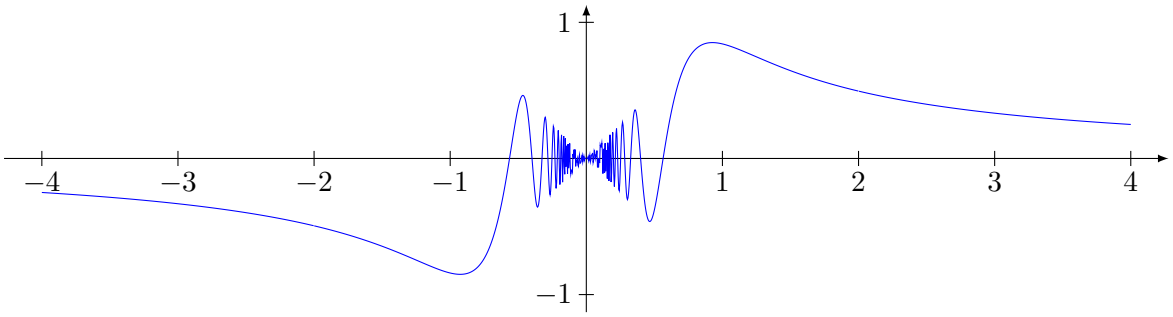


FIGURE 19.12 – La fonction de l'exemple 19.208.

19.13.2 Méthode du développement asymptotique

Nous savons que nous pouvons développer certaines fonctions en série grâce au développement de Taylor (théorème 13.359). Lorsque nous avons une limite à calculer, nous pouvons remplacer certaines parties de la fonction à traiter par la formule (13.948b). Cela est très utile pour comparer des fonctions trigonométrique à des polynômes.

Exemple 19.209
Nous calculons la limite

$$\lim_{x\to 0} \frac{\sin(x)}{x} \tag{19.729}$$

à l'aide d'un tout petit développement en série. Pour information, la proposition 19.211 nous fera la même limite sans développements.

Une manière de la prouver est d'écrire

$$\sin(x) = x + h(x) \tag{19.730}$$

avec $h \in o(x)$, c'est-à-dire $\lim_{x\to 0} h(x)/x = 0$. Alors nous avons

$$\lim_{x\to 0} \frac{\sin(x)}{x} = \lim_{x\to 0} \frac{x + h(x)}{x} = \lim_{x\to 0} \frac{x}{x} + \lim_{x\to 0} \frac{h(x)}{x} = 1. \tag{19.731}$$

△

L'utilisation de la proposition 13.395 permet d'utiliser cette technique dans le cadre de limites à plusieurs variables. Reprenons l'exemple 19.209 un tout petit peu modifié :

Exemple 19.210
Soit à calculer $\lim_{(x,y)\to(0,0)} f(x,y)$ où

$$f(x,y) = \frac{\sin(xy)}{xy}. \tag{19.732}$$

La première chose à faire est de voir f comme la composée de fonctions $f = f_1 \circ f_2$ avec

$$\begin{aligned} f_1 \colon \mathbb{R} &\to \mathbb{R} \\ t &\mapsto \frac{\sin(t)}{t} \end{aligned} \tag{19.733}$$

et

$$\begin{aligned} f_2 \colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto xy. \end{aligned} \tag{19.734}$$

Étant donné que $\lim_{(x,y)\to(0,0)} f_2(x,y) = 0$, nous avons $\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{t\to 0} f_1(t) = 1$. △

La proposition suivante reprend la limite de l'exemple 19.209.

Proposition 19.211.
Nous avons la limite

$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1. \tag{19.735}$$

Démonstration. On commence par observer que la fonction $g(x) = \frac{\sin(x)}{x}$ est un rapport entre deux fonctions impaires et est donc une fonction paire. Nous pouvons alors nous réduire à considérer le cas où x est positif. La première étape de cette preuve nous dit que $g(x) \leqslant 1$ pour tout $x \in \mathbb{R}^{+,*}$.

Nous voulons étudier le comportement de g dans un voisinage de 0. Nous pouvons alors supposer que x soit inférieur à $\pi/2$. Soit $D = (1, \tan(x))$. On voit très bien dans le dessin que l'aire du triangle de sommets O, D et C est supérieure à l'aire du secteur circulaire de sommets O, A et C. Ces deux aires peuvent être calculées très facilement et nous obtenons

$$\frac{\sin(x)}{2\cos(x)} \geqslant \frac{x}{2}.$$

À partir de cette dernière inégalité nous pouvons écrire

$$1 \geqslant \frac{\sin(x)}{x} \geqslant \cos(x).$$

En prenant la limite lorsque x tend vers 0 dans les trois membres de l'inégalité la règle de l'étau nous permet d'obtenir la limite remarquable (19.735). □

Exemple 19.212
Les dérivées partielles de la fonction $f(x,y) = xy^3 + \sin y$ au point $(0,\pi)$ sont

$$\partial_x f(0,\pi) = \frac{\partial f}{\partial x}(0,\pi) = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{(t\pi^3 + \sin\pi) - (\sin\pi)}{t} = \pi^3,$$

$$\partial_y f(0,\pi) = \frac{\partial f}{\partial y}(0,\pi) = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{0(\pi+t)^3 + \sin(t+\pi) - 0\cdot\pi^3}{t} = \cos\pi = -1,$$

△

19.14 Changement de variables dans une intégrale

Le théorème 15.252 manque un peu d'exemples. Nous allons en voir quelques-uns maintenant.

19.14.1 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

$$\int_E f = \int_0^{2\pi}\int_0^1 f(r\cos\theta, r\sin\theta) dr d\theta. \tag{19.736}$$

Hélas, non ; la vie n'est pas aussi simple.

Définition 19.213.
*Un **difféomorphisme** est une application $g: A \to B$ telle que g et $g^{-1}: B \to A$ soient de classe C^1.*

Théorème 19.214.
Soit $g: A \to B$ un difféomorphisme. Soient $F \subset B$ un ensemble mesurable et borné et $f: F \to \mathbb{R}$ une fonction bornée et intégrable. Supposons que $g^{-1}(F)$ soit borné et que Jg soit borné sur $g^{-1}(F)$. Alors

$$\int_F f(x)dy = \int_{g^{-1}(F)f\big(g(x)\big)} |Jg(x)|dx \tag{19.737}$$

Pour rappel, Jg est le déterminant de la matrice jacobienne (aucun lien de parenté) donnée par

$$Jg = \det \begin{pmatrix} \partial_x g_1 & \partial_y g_1 \\ \partial_x g_2 & \partial_t g_2 \end{pmatrix}. \tag{19.738}$$

Comme dans les intégrales simples, il y a souvent moyen de trouver un changement de variables qui simplifie les expressions. Le domaine $E = \{(x, y) \in \mathbb{R}^2 \text{ tel que } x^2 + y^2 < 1\}$ par exemple 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

$$\int_E f = \int_0^{2\pi} \int_0^1 f(r\cos\theta, r\sin\theta)drd\theta. \tag{19.739}$$

Hélas ce n'est pas le cas. Il faut tenir compte du fait que le changement de base dilate ou contracte certaines surfaces.

Soit $\varphi: D_1 \subset \mathbb{R}^2 \to D_2 \subset \mathbb{R}^2$ 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

$$\varphi(u, v) = \begin{pmatrix} x(u, v) \\ y(u, v) \end{pmatrix} \tag{19.740}$$

avec $(u, v) \in D_1$.

Théorème 19.215.
Soit une fonction continue $f: D_2 \to \mathbb{R}$. Alors

$$\int_{\varphi(D_1)} f(x, y)dxdy = \int_{D_1} f\big(x(u, v), y(u, v)\big)|J_\varphi(u, v)|dudv \tag{19.741}$$

où J_φ est le Jacobien de φ c'est-à-dire

$$J_\varphi(u, v) = \det \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial u}{\partial v} \end{pmatrix}. \tag{19.742}$$

Ne pas oublier de prendre la valeur absolue lorsqu'on utilise le Jacobien dans un changement de variables.

19.14.2 Coordonnées polaires

Exemple 19.216
Calculons la surface du disque D de rayon R. Nous devons calculer

$$\int_D dxdy. \tag{19.743}$$

Pour passer au polaires, nous savons que le disque est décrit par

$$D = \{(r, \theta) \text{ tel que } 0 \leqslant r \leqslant R, 0 \leqslant \theta \leqslant 2\pi\}. \tag{19.744}$$

Nous avons donc

$$\int_D dxdy = \int_D r\,drd\theta = \int_0^{2\pi}\int_0^R r\,drd\theta = 2\pi\int_0^R r\,dr = \pi R^2. \tag{19.745}$$

△

Exemple 19.217
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 19.13. 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 : $y^2 - y = (y - \frac{1}{2})^2 - \frac{1}{4}$.

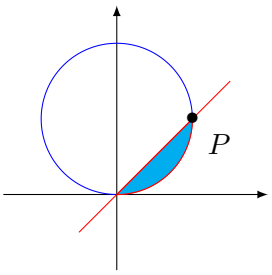


FIGURE 19.13 – Passage en polaire pour intégrer sur un morceau de cercle.

Le passage en polaire transforme les équations du bord du domaine en

$$\begin{aligned}\cos(\theta) &= \sin(\theta)\\ r^2 &= r\sin(\theta).\end{aligned} \tag{19.746}$$

L'angle θ parcours donc $]0,\pi/4[$, et le rayon, pour chacun de ces θ parcours $]0,\sin(\theta)[$. La fonction à intégrer se note maintenant $f(r,\theta) = \sqrt{1-r^2}$. Donc l'intégrale à calculer est

$$\int_0^{\pi/4}\left(\int_0^{\sin(\theta)}\sqrt{1-r^2}r\,rd\right). \tag{19.747}$$

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 « $dxdy$ » devient « $r\,drd\theta$ ».

Exemple 19.218
On veut évaluer l'intégrale de la fonction $f(x,y) = x^2+y^2$ sur la région V suivante :

$$V = \{(x,y)\in\mathbb{R}^2\,|\,x^2+y^2\leqslant 1,\ x>0,\ y>0\}.$$

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_ϕ est

$$J_\phi(r,\theta) = \begin{vmatrix}\cos\theta & \sin\theta\\ -r\sin\theta & r\cos\theta\end{vmatrix} = r, \tag{19.748}$$

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

$$\int_V f(x,y)\,dV = \int_0^{\pi/2}\int_0^1 r^3 dr\,d\theta = \int_0^{\pi/2}\frac{1}{4}\,d\theta = \frac{\pi}{8}.$$

△

Exemple 19.219
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 19.14. 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 : $y^2 - y = (y-\frac{1}{2})^2 - \frac{1}{4}$.

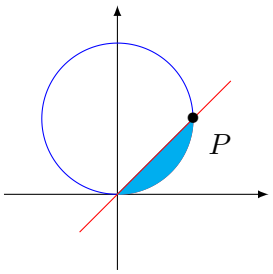


FIGURE 19.14 – Passage en polaire pour intégrer sur un morceau de cercle.

Le passage en polaire transforme les équations du bord du domaine en

$$\begin{aligned}\cos(\theta) &= \sin(\theta)\\ r^2 &= r\sin(\theta).\end{aligned} \tag{19.749}$$

L'angle θ parcours donc $]0,\pi/4[$, et le rayon, pour chacun de ces θ parcours $]0,\sin(\theta)[$. La fonction à intégrer se note maintenant $f(r,\theta) = \sqrt{1-r^2}$. Donc l'intégrale à calculer est

$$\int_0^{\pi/4}\left(\int_0^{\sin(\theta)}\sqrt{1-r^2}r\,rd\right). \tag{19.750}$$

Remarquez la présence d'un r supplémentaire pour le jacobien.

Notez que les coordonnées du point P sont $(1,1)$. △

19.14.3 Coordonnées cylindriques

Exemple 19.220
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 $(0,1/2)$

$$A = \{(x,y,z)\in\mathbb{R}^3\,|\,x^2+y^2+z^1\leqslant 1\}\cap\{(x,y,z)\in\mathbb{R}^3\,|\,x^2+(y-1/2)^2\leqslant 1/4\}.$$

On peut décrire A en coordonnées cylindriques

$$\begin{aligned}A = \Big\{(r,\theta,z)\in{}&]0,+\infty[\times[-\pi,\pi[\times\mathbb{R}\,|\\ &-\pi/2<\theta<\pi,\ 0<r\leqslant\sin\theta,\ -\sqrt{1-r^2}\leqslant z\leqslant\sqrt{1-r^2}\Big\}.\end{aligned} \tag{19.751}$$

Le jacobien de ce changement de variables, J_{cyl}, est

$$J_{cyl}(r,\theta),z = \begin{vmatrix}\cos\theta & \sin\theta & 0\\ -r\sin\theta & r\cos\theta & 0\\ 0 & 0 & \end{vmatrix} = r, \tag{19.752}$$

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 19.221(Volume d'un solide de révolution)
Soit $g : [a,b] \to \mathbb{R}_+$ une fonction continue et positive. On dit que le solide A décrit par

$$A = \left\{(x,y,z) \in \mathbb{R}^3 \,|\, z \in [a,b],\ \sqrt{x^2+y^2} \leqslant g^2(z)\right\}$$

est un solide de révolution. Afin de calculer son volume, on peut décrire A en coordonnées cylindriques,

$$A = \left\{(r,\theta,z) \in]0,+\infty[\times[-\pi,\pi[\times\mathbb{R} \,|\, a \leqslant z \leqslant b,\ 0 < r^2 \leqslant g^2(z)\right\}.$$

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. △

19.14.3.1 Coordonnées sphériques

Le calcul est un peu plus long :

$$\begin{aligned} J(\rho,\theta,\varphi) &= \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{vmatrix} \\ &= \begin{vmatrix} \sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & \rho\cos\theta\sin\varphi & -\rho\sin\theta\cos\varphi \\ \cos\theta & -\rho\sin\theta & 0 \end{vmatrix} \\ &= \rho^2\sin\theta. \end{aligned} \tag{19.753}$$

Donc

$$dx\,dy\,dz = \rho^2\sin(\theta)\,d\rho\,d\theta\,d\varphi. \tag{19.754}$$

19.14.4 Coordonnées sphériques

Exemple 19.222
On veut calculer le volume du cornet de glace A

$$A = \left\{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in \mathbb{S}^2,\ \sqrt{x^2+y^2} \leqslant z \leqslant \sqrt{1-x^2-y^2}\right\}.$$

On peut décrire A en coordonnées sphériques.

$$A = \{(\rho,\theta,\phi) \in]0,+\infty[\times[-\pi,\pi[\times[0,\pi[\,|\, 0 < \phi \leqslant \pi/4,\ 0 < \rho \leqslant 1\}.$$

Le jacobien de ce changement de variables J_{sph} est

$$J_{sph}(\rho,\theta,\phi) = \begin{vmatrix} \cos\theta\sin\phi & \sin\theta\sin\phi & \cos\phi \\ -\rho\sin\theta\sin\phi & \rho\cos\theta\sin\phi & 0 \\ \rho\cos\theta\cos\phi & \rho\sin\theta\cos\phi & -\rho\sin\phi \end{vmatrix} = \rho^2\sin\phi, \tag{19.755}$$

Le volume de A est donc

$$\int_{\mathbb{R}^3} \chi_A(x,y,z)\,dV = \int_{-\pi}^{\pi}\int_0^{\pi/4}\int_0^1 \rho^2\sin\phi\,d\rho\,d\phi\,d\theta = \frac{2\pi}{3}\left(1-\frac{1}{\sqrt{2}}\right).$$

△

Exemple 19.223(Une petite faute à ne pas faire)
Si nous voulons calculer le volume de la sphère de rayon R, nous écrivons donc

$$\int_0^R dr\int_0^{2\pi} d\theta\int_0^{\pi} r^2\sin(\phi)d\phi = 4\pi R = \frac{4}{3}\pi R^3. \tag{19.756}$$

Ici, la valeur absolue n'est pas importante parce que lorsque $\phi\in\,]0,\pi,[$, le sinus de ϕ est positif.

Des petits malins pourraient remarquer que le changement de variable (19.696) est encore une paramétrisation de $\mathbb{R}^3$ si on intervertit le domaine des angles :

$$\begin{aligned}\theta&\colon 0\to\pi\\ \phi&\colon 0\to 2\pi,\end{aligned} \tag{19.757}$$

alors nous paramétrons encore parfaitement bien la sphère, mais hélas

$$\int_0^R dr\int_0^{\pi} d\theta\int_0^{2\pi} r^2\sin(\phi)d\phi = 0. \tag{19.758}$$

Pourquoi ces « nouvelles » coordonnées sphériques sont-elles mauvaises ? Il y a que quand l'angle ϕ parcours $]0,2\pi[$, 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 ϕ entre 0 et π, puis $-r^2\sin(\phi)$ pour ϕ entre π et 2π. Donc l'intégrale (19.758) n'est pas correcte. Il faut la remplacer par

$$\int_0^R dr\int_0^{\pi} d\theta\int_0^{\pi} r^2\sin(\phi)d\phi - \int_0^R dr\int_0^{\pi} d\theta\int_{\pi}^{2\pi} r^2\sin(\phi)d\phi = \frac{4}{3}\pi R^3 \tag{19.759}$$

△

19.14.5 Un autre système utile

Un changement de variables que l'on voit assez souvent est

$$\begin{cases} u = x+y & \text{(19.760a)}\\ v = x-y. & \text{(19.760b)}\end{cases}$$

Afin de calculer son jacobien, il faut d'abord exprimer x et y en fonctions de u et v :

$$\begin{cases} x = (u+v)/2 & \text{(19.761a)}\\ y = (u-v)/2. & \text{(19.761b)}\end{cases}$$

La matrice jacobienne est

$$\begin{pmatrix}\frac{\partial x}{\partial u} & \frac{\partial x}{\partial v}\\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v}\end{pmatrix} = \begin{pmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{pmatrix}. \tag{19.762}$$

Le déterminant vaut $-\frac{1}{2}$. Nous avons donc

$$dxdy = \frac{1}{2}dudv. \tag{19.763}$$

Nous insistons sur le fait que c'est $\frac{1}{2}$ et non $-\frac{1}{2}$ qui intervient parce que que la formule du changement de variable demande d'introduire la *valeur absolue* du jacobien.

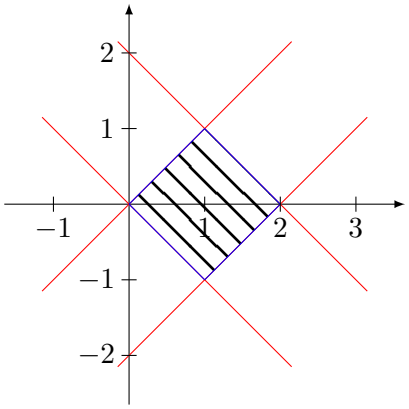


FIGURE 19.15 – Un domaine qui s'écrit étonnament bien avec un bon changement de coordonnées.

Exemple 19.224
Calculer l'intégrale de la fonction $f(x,y) = x^2 - y^2$ sur le domaine représenté sur la figure 19.15.

Les droites qui délimitent le domaine d'intégration sont

$$\begin{aligned} y &= -x + 2 \\ y &= x - 2 \\ y &= x \\ y &= -x \end{aligned} \tag{19.764}$$

Le domaine est donc donné par les équations

$$\begin{cases} y + x < 2 & \text{(19.765a)} \\ y - x > -2 & \text{(19.765b)} \\ y - x < 0 & \text{(19.765c)} \\ y + x > 0. & \text{(19.765d)} \end{cases}$$

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

$$f(u,v) = uv. \tag{19.766}$$

L'intégrale à calculer est simplement

$$\int_0^2 \int_0^2 uv\, du dv = \int_0^2 u\, du \left[\frac{v^2}{2}\right]_0^2 = 2\int_0^2 u\, du = 4. \tag{19.767}$$

△

19.15 Table de caractères du groupe diédral

Cette section vient de [43] ; nous avons comme but d'établir la table des caractères des représentations complexes du groupe diédral D_n.

19.15.1 Représentations de dimension un

Nous nous occupons des représentations de D_n sur $\mathbb{C}$. Les applications linéaires $\mathbb{C} \to \mathbb{C}$ sont seulement les multiplications par des nombres complexes. Nous cherchons donc $\psi\colon D_n \to \mathbb{C}^*$.

Nous savons que D_n est généré[59] par s et r. Vu que $s^2 = 1$, nous avons

$$\psi(s)^2 = \psi(s^2) = \psi(1) = 1, \tag{19.768}$$

donc $\psi(s) \in \{-1, 1\}$. Nous savons aussi que $srsr = 1$, donc

$$\psi(s)^2\psi(r)^2 = 1, \tag{19.769}$$

ce qui donne $\psi(r) \in \{-1, 1\}$.

Nous avons donc quatre représentations de dimension un données par

	$\psi(r) = 1$	$\psi(r) = -1$
$\psi(s) = 1$	ρ^{++}	ρ^{+-}
$\psi(s) = -1$	ρ^{-+}	ρ^{--}

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
χ^{++}	1	1
χ^{+-}	$(-1)^k$	$(-1)^k$
χ^{-+}	1	-1
χ^{--}	$(-1)^k$	$(-1)^{k+1}$

Étant donné qu'ils sont tous différents, ce sont des représentations deux à deux non équivalentes, lemme 17.15.

19.15.2 Représentations de dimension deux

Nous cherchons maintenant les représentations $\rho\colon D_n \to \text{End}(\mathbb{C}^2)$. Ici nous supposons connue la liste des éléments de D_n donnée par le corollaire 19.95. Soit $\omega = e^{2i\pi/n}$ et $h \in \mathbb{Z}$; nous considérons la représentation $\rho^{(h)}$ de D_n définie par

$$\rho^{(h)}(r^k) = \begin{pmatrix} \omega^{hk} & 0 \\ 0 & \omega^{-hk} \end{pmatrix} \tag{19.770a}$$

$$\rho^{(h)}(sr^k) = \begin{pmatrix} 0 & \omega^{-hk} \\ \omega^{hk} & 0 \end{pmatrix}. \tag{19.770b}$$

Cela donne bien $\rho^{(h)}$ sur tous les éléments de D_n par la proposition 19.94. 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 $0 \leqslant h \leqslant \frac{n}{2}$.

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 s'il vérifie les systèmes d'équations

$$\begin{cases} ax = \lambda x & \text{(19.771a)} \\ \frac{1}{a}y = \lambda y & \text{(19.771b)} \end{cases}$$

et

$$\begin{cases} \frac{1}{a}y = \mu x & \text{(19.772a)} \\ ax = \mu y & \text{(19.772b)} \end{cases}$$

59. Voir proposition 19.94 et tout ce qui suit.

avec λ et μ 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

$$\rho^{(0)}(r^k) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \rho^{(0)}(sr^k) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{19.773}$$

donc le caractère de cette représentation est $\chi^{(0)}(r^k) = 2$ et $\chi^{(0)}(sr^k) = 0$. Donc nous avons

$$\chi^{(0)} = \chi^{++} + \chi^{-+}. \tag{19.774}$$

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 :

$$S(r) = (\rho^{++} \oplus \rho^{-+})(r^k) = \begin{pmatrix} \rho^{++}(r^k) & 0 \\ 0 & \rho^{-+}(r^k) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{19.775a}$$

$$S(sr^k) = (\rho^{++} \oplus \rho^{-+})(sr^k) = \begin{pmatrix} \rho^{++}(sr^k) & 0 \\ 0 & \rho^{-+}(sr^k) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{19.775b}$$

(19.775c)

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

$$T \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{19.776}$$

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 17.22(1) et dire que les représentations sont équivalentes parce que les caractères sont égaux.

Troisième méthode Utiliser le théorème 17.22(2) et calculer $\langle \chi^{(0)}, \chi^{(0)} \rangle$:

$$\langle \chi^{(0)}, \chi^{(0)} \rangle = \frac{1}{|D_n|} \sum_{g \in D_n} |\chi^{(0)}(g)|^2 \tag{19.777a}$$

$$= \frac{1}{2n}\big(4 + 0 + 4(n-1)\big) \tag{19.777b}$$

$$= 2. \tag{19.777c}$$

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 (19.771) et (19.772) 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 a 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

$$\rho^{(n/2)}(r^k) = \begin{pmatrix} (-1)^k & 0 \\ 0 & (-1)^k \end{pmatrix} \quad \rho^{(n/2)}(sr^k) = \begin{pmatrix} 0 & (-1)^k \\ (-1)^k & 0 \end{pmatrix} , \tag{19.778}$$

et donc

$$\chi^{(n/2)}(r^k) = 2(-1)^k \tag{19.779a}$$
$$\chi^{(n/2)}(sr^k) = 0. \tag{19.779b}$$

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 $T \in \mathrm{GL}(2, \mathbb{C})$ 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

19.15.3 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

$$\frac{n}{2} + 3 \tag{19.780}$$

représentations irréductibles modulo équivalence. Cela fait le compte en vertu des classes de conjugaisons listées en 19.9.1.5. Pour rappel, le nombre de représentations non équivalentes est égal au nombre de classes de conjugaison par le corollaire 17.25. 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.

19.15.4 Le compte pour n impair

Nous avions fait mention plus haut du fait que si ψ est une représentation de dimension 1, le nombre $\psi(r)$ devait être une racine n^{e} de l'unité. Donc en dimension 1 nous avons seulement les représentations ρ^{++} et ρ^{-+}. Pour celles de dimension 2, nous en avons $\frac{n-1}{2}$. En tout nous avons donc

$$\frac{n+3}{2} \tag{19.781}$$

représentations irréductibles modulo équivalence. Cela fait le compte en vertu des classes de conjugaisons listées en 19.9.1.6.

Chapitre 20

Corps finis, racines de l'unité

20.1 Le groupe des racines de l'unité dans les nombres complexes

20.1.1 Le groupe

Définition 20.1.
Une ***racine*** n^e ***de l'unité*** *dans un anneau est une racine du polynôme* $X^n - 1$.

Lemme-définition 20.2.
L'ensemble

$$\mathrm{U}_n = \{e^{2i\pi k/n} \text{ tel que } k = 0, \dots, n-1\} \tag{20.1}$$

est un groupe pour la multiplication.

Ses éléments forment exactement l'ensemble des racines de l'unité dans $\mathbb{C}$.

Démonstration. Il est vite vu que tous les éléments de U_n sont des racines de l'unité parce que

$$\left(e^{2i\pi k/n}\right)^n = e^{2i\pi k} = 1, \tag{20.2}$$

entre autres à cause du lemme 19.25.

Cela nous donne déjà n racines pour $X^n - 1$ dans $\mathbb{C}$. Le théorème 6.99 nous indique qu'il ne peut pas y en avoir plus. □

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 20.18.

Lemme 20.3.
L'ensemble U_n *est un groupe cyclique*[1] *d'ordre* n *généré par* $\xi = e^{2i\pi/n}$.

Démonstration. Il y a les trois propriétés à vérifier pour que ce soit un groupe.

Neutre Le nombre 1 est une racine de l'unité.

Inverse Si $\omega \in U_n$ alors $\omega^n = 1$ et donc $\omega\omega^{n-1} = 1$, ce qui signifie que ω^{n-1} est un inverse de ω.
Il reste à voir que $\omega^{n-1} \in U_n$. En effet $\left(\omega^{n-1}\right)^n = (\omega^n)^{n-1} = 1^{n-1} = 1$.

Associativité Cas particulier de l'associativité dans $\mathbb{C}$.

Le fait que ce soit un groupe cyclique contenant n éléments est la définition. □

Le lemme suivant donne les autres générateurs.

Lemme 20.4.
Le nombre ξ^a *est un générateur de* U_n *si et seulement si* $\gcd(a, n) = 1$.

1. Définition 2.11.

Démonstration. Si $\text{pgcd}(a,n)=1$ alors le théorème de Bézout 3.13 nous fournit des entiers u et v tels que $ua+vn=1$. Alors nous avons

$$e^{2i\pi/n}=e^{2(ua+vn)i\pi/n}=(e^{2ai\pi/n})^u, \tag{20.3}$$

ce qui signifie que ξ est dans le groupe engendré par ξ^a, et par conséquent tout U_n est engendré.

Pour l'implication inverse, nous utilisons Bézout dans le sens inverse. Soit ξ^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 $\text{pgcd}(a,n)=1$. □

Exemple 20.5
Une conséquence tout à fait extraordinaire de ce lemme est que 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[331]. Ce fait est connu des pianistes [2] depuis des siècles. △

Proposition 20.6 (Intersection par deux)**.**
Les ensembles U_α et U_β ont une intersection réduite à $\{1\}$ si et seulement si α et β sont premiers entre eux.

Démonstration. Nous rappelons qu'une racine α^{e}de l'unité peut s'écrire sous la forme $e^{2i\pi k/\alpha}$ avec $0\leqslant k<\alpha$.

Sens direct Par contraposée, nous supposons que α et β 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_α et U_β nous devons trouver une valeur de $0<k<\alpha$ telle que

$$(e^{2i\pi k/\alpha})^\beta=e^{2i\pi k\beta/\alpha}=1, \tag{20.4}$$

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 $k=\alpha'<\alpha$.

Sens réciproque Supposons maintenant que α et β soient premiers entre eux. Soit $z\in U_\alpha\cap U_\beta$. Le fait que z soit une racine α^{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 β^{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

$$k\beta=l\alpha. \tag{20.5}$$

Si $k>0$, comme le nombre α divise $k\beta$, cela conduirait via le lemme de Gauss 3.21 à dire que α divise k. Mais α ne peut pas diviser k parce que nous avions supposé que k était strictement plus petit que α. Donc $k=0$ et $z=1$.

□

Proposition 20.7 (Intersection : le cas général[1])**.**
Soient des entiers positifs $\alpha_1,\dots,\alpha_p$. Nous avons

$$\bigcup_{i=1}^{p}U_{\alpha_i}=\{1\} \tag{20.6}$$

si et seulement si $\text{pgcd}(\alpha_1,\dots,\alpha_p)=1$ *(c'est-à-dire que les α_i sont premiers dans leur ensemble).*

Démonstration. Nous le décomposons les α_i en facteurs premiers [3] de la façon suivante : $\alpha_i=\prod_{k\in\mathbb{N}}p_k^{\alpha_i^{(k)}}$ où les p_k sont les nombres premiers.

2. Même ceux qui ignorent le théorème de Bézout.
3. Théorème 3.23.

Caractérisation par une décomposition en facteurs premiers Les éléments z différents de 1 dans U_{α_1} s'écrivent sous la forme

$$z = e^{2i\pi k/\alpha_1} \tag{20.7}$$

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 $\left(e^{2i\pi k/\alpha_1}\right)^{\alpha_i} = 1$, c'est-à-dire que $\alpha_i k/\alpha_1$ est entier, donc que α_1 divise $k\alpha_i$. Par conséquent il existera un élément différent de 1 dans l'intersection des U_{α_i} si et seulement s'il existe un entier k strictement compris entre 0 et α_1 pour lequel α_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

$$\alpha_1^{(l)} - \alpha_i^{(l)} \tag{20.8}$$

pour tout l.

Sens direct L'hypothèse $\operatorname{pgcd}(\alpha_1, \dots, \alpha_p) \neq 1$ 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 α_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 α_1.

Sens réciproque Si $\operatorname{pgcd}(\alpha_1, \dots, \alpha_p) = 1$ 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 $k \geqslant \alpha_1$, ce qui est impossible.

□

Définition 20.8.
Les générateurs de U_n *sont les* ***racines primitives***[4] *de l'unité dans* $\mathbb{C}$. *Nous nommons* Δ_n *leur ensemble :*

$$\Delta_n = \{e^{2ki\pi/n} \text{ tel que } 0 \leqslant k \leqslant n-1, \operatorname{pgcd}(k,n) = 1\}. \tag{20.9}$$

Nous avons par exemple

$$\Delta_1 = \{1\} \tag{20.10a}$$
$$\Delta_2 = \{e^{\pi i}\} \tag{20.10b}$$
$$\Delta_4 = \{e^{\pi i/2}, e^{3\pi i/2}\}. \tag{20.10c}$$

Notons que $1 \in \Delta_d$ seulement avec $d = 1$.

20.1.2 Fonction indicatrice d'Euler (première partie)

Nous introduisons ici la fonction indicatrice d'Euler et ses liens basiques avec les racines de l'unité. Pour les propriétés plus avancées, voir 20.1.5.

20.1.3 Introduction par les racines de l'unité

Définition 20.9.
La fonction φ *donnée par*

$$\varphi(n) = \operatorname{Card}(\Delta_n) \tag{20.11}$$

*est l'****indicatrice d'Euler****.*

Si p est un nombre premier, alors $\varphi(p) = p - 1$.

4. 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 20.57.

Lemme 20.10.
Nous avons

$$\mathrm{U}_n = \bigcup_{d\,|\,n} \Delta_d \tag{20.12}$$

et l'union est disjointe. Nous avons aussi la formule

$$n = \sum_{d\,|\,n} \varphi(d). \tag{20.13}$$

Démonstration. À l'application $x \mapsto e^{2i\pi x}$ près, nous pouvons considérer

$$\Delta_d = \{\frac{k}{d} \text{ tel que } k = 0, \dots, d-1, \mathrm{pgcd}(k,d) = 1\}, \tag{20.14}$$

c'est-à-dire l'ensemble des fractions irréductibles dont le dénominateur est d. L'union des Δ_d sera donc disjointe.

Toujours à l'application $x \mapsto e^{2i\pi x}$ près, le groupe U_n est donné par

$$\mathrm{U}_n = \{\frac{k}{n} \text{ tel que } k = 0, \dots, n-1\}. \tag{20.15}$$

L'égalité (20.12) revient maintenant à dire que toute fraction de la forme $\frac{k}{n}$ s'écrit de façon irréductible avec un dénominateur qui divise n.

La relation (20.13) consiste à prendre le cardinal des deux côtés de (20.12). Nous avons $\mathrm{Card}(\mathrm{U}_n) = n$ et l'union étant disjointe, à droite nous avons la somme des cardinaux.

Pour chaque diviseur d de n nous considérons l'ensemble

$$\Phi_n(d) = \{m \in \mathbb{N} \text{ tel que } \mathrm{pgcd}(m,n) = \frac{n}{d}\}. \tag{20.16}$$

Étant donné que tous les entiers entre 1 et n ont un pgcd avec n qui est automatiquement un quotient de n nous avons

$$\{1, \dots, n\} = \bigcup_{d\,|\,n} \Phi_n(d) \tag{20.17}$$

où l'union est disjointe. Par ailleurs nous savons que si $\mathrm{pgcd}(a,b) = 1$, alors $\mathrm{pgcd}(ka, kb) = k$. Donc si $m \in \Delta_d$, alors $m \cdot \frac{n}{d}$ appartient à $\Phi_n(d)$. En d'autres termes, $a \mapsto \frac{n}{d}a$ est une bijection entre Δ_d et $\Phi_n(d)$.

Nous avons donc $\mathrm{Card}(\Phi_n(d)) = \mathrm{Card}(\Delta_d) = \varphi(d)$ et finalement

$$\mathrm{Card}\{1, \dots, n\} = \sum_{d\,|\,n} \mathrm{Card}(\Phi_n(d)) = \sum_{d\,|\,n} \varphi(d). \tag{20.18}$$

□

Lemme 20.11.
Si p est un nombre premier, alors $\varphi(p^n) = p^n - p^{n-1}$.

Démonstration. Les éléments de $\{1, \dots, p^n\}$ qui ont un pgcd différent de 1 avec p^n sont des nombres qui s'écrivent sous la forme qp avec $q \leqslant p^{n-1}$ [5]. Il y a évidemment p^{n-1} tels nombres.

Par conséquent le cardinal de P_{p^n} est $\varphi(p^n) = p^n - p^{n-1}$. □

Problèmes et choses à faire

P_n n'a pas été défini.

Définition proposée (et vue par après) : $P_n = \{m \in \mathbb{N} \text{ tel que } \mathrm{pgcd}(m,n) = 1\}$. À mettre donc en lien avec Δ_d.

5. Corollaire 3.25.

20.1.4 Générateurs

Proposition 20.12.
Soit $n \in \mathbb{N}^*$ et le groupe (additif) $\mathbb{Z}/n\mathbb{Z}$. L'élément $[x]_n$ est un générateur de $\mathbb{Z}/n\mathbb{Z}$ si et seulement si $x \in P_n$. En particulier $\mathbb{Z}/n\mathbb{Z}$ est un groupe contenant $\varphi(n)$ générateurs.

Démonstration. Nous avons $\text{gr}\big([1]_n\big) = \mathbb{Z}/n\mathbb{Z}$. L'élément $[x]_n$ sera générateur si et seulement s'il génère $[1]_n$, c'est-à-dire s'il existe u tel que $u[x]_n = [1]_n$. Cette dernière égalité étant une égalité de classes dans $\mathbb{Z}/n\mathbb{Z}$, elle sera vraie si et seulement s'il existe v tel que

$$ux + vn = 1. \tag{20.19}$$

Cela signifie entre autres que[6] $x\mathbb{Z} + n\mathbb{Z} = \mathbb{Z}$, et aussi que $\text{pgcd}(x, n) = 1$ par le théorème de Bézout 3.13, et donc que $x \in P_n$. □

Corollaire 20.13.
Un groupe monogène d'ordre n possède $\varphi(n)$ générateur où φ est la fonction indicatrice d'Euler définie en 20.9.

Démonstration. Le théorème 5.16 nous dit qu'un groupe monogène d'ordre n est isomorphe à $\mathbb{Z}/n\mathbb{Z}$. La proposition 20.12 nous indique que $\mathbb{Z}/n\mathbb{Z}$ possède $\varphi(n)$ générateurs. □

20.1.5 Fonction indicatrice d'Euler (propriétés)

Corollaire 20.14.
L'indicatrice d'Euler est multiplicative : si p est premier avec q, alors $\varphi(pq) = \varphi(p)\varphi(q)$. De plus si p et q sont premiers entre eux,

$$\varphi(pq) = (p-1)(q-1). \tag{20.20}$$

Démonstration. Nous savons que si p et q sont premiers entre eux, alors le théorème 5.25 nous donne l'isomorphisme de groupe

$$(\mathbb{Z}/pq\mathbb{Z}, +) \simeq (\mathbb{Z}/p\mathbb{Z}, +) \times (\mathbb{Z}/q\mathbb{Z}, +). \tag{20.21}$$

Un élément (x, y) est générateur du produit si et seulement si x est générateur de $\mathbb{Z}/p\mathbb{Z}$ et y est générateur de $\mathbb{Z}/q\mathbb{Z}$. Par la proposition 20.12, il y a $\varphi(p)\varphi(q)$ tels éléments. Par ailleurs le nombre de générateurs de $\mathbb{Z}/pq\mathbb{Z}$ est $\varphi(pq)$, d'où l'égalité.

Si p est premier, nous avons $\varphi(p) = p - 1$ parce que tous les entiers de $\{1, \dots, p-1\}$ sont premiers avec p. □

20.1.6 Décomposition en fractions simples

Il est possible de décomposer une fraction rationnelle en fractions dites « simples ». Si $|z| < 1$ nous avons par exemple la décomposition

$$\frac{1}{1-z^r} = \sum_{\omega \in U_r} \frac{A_\omega}{\omega - z} \tag{20.22}$$

où U_r est le groupe des racines r^{e} de l'unité définit en (20.1). Les nombres A_ω peuvent alors être déterminés en effectuant la somme. Le dénominateur commun sera $1 - z^r$ tandis que les A_ω sont déterminés en égalant le numérateur à 1.

Exemple 20.15
Pour décomposer la fraction $\frac{1}{1-x^2}$ nous savons que les racines sont ± 1. Donc nous écrivons

$$\frac{1}{1-x^2} = \frac{A}{1-x} + \frac{B}{1+x}. \tag{20.23}$$

6. Corollaire 3.14

Nous trouvons les valeurs de A et B en effectuant la somme :

$$\frac{A(1+x)+B(1-x)}{1-x^2}=\frac{A+B+(A-B)x}{1-x^2}. \tag{20.24}$$

Les coefficients A et B doivent donc vérifier $A+B=1$ et $A-B=0$. Au final,

$$\frac{1}{1-x^2}=\frac{1}{2(1-x)}+\frac{1}{2(1+x)}. \tag{20.25}$$

△

20.2 Chiffrement RSA

Ce passage sur RSA provient en bonne partie de la page de Wikipédia[332].

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.

20.2.1 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 corollaire 20.14, 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) [7].

Bob envoie la paire (n,e) à Alice.

Remarque 20.16.
Ici nous ne supposons pas que la communication soit sure. Une tierce personne peut intercepter le message. D'ailleurs en principe les gens publient leurs clef publique sur leurs sites, voire sur des sites dédiés. Le problème de l'identification reste à résoudre à l'ancienne.

20.2.2 Chiffrement

Nous chiffrons 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

$$C=M^e \mod n. \tag{20.26}$$

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.

20.2.3 Déchiffrement

Nous allons montrer que $M=C^d \mod n$, et donc que Bob, connaissant (n,d), peut déchiffrer. D'abord

$$C^d=(M^e)^d=M^{ed}, \tag{20.27}$$

mais nous savons qu'il existe k tel que

$$ed=1+k\varphi(n)=1+k(p-1)(q-1). \tag{20.28}$$

7. Le fait que e soit public et d soit privé est une convention. e comme *encryption* et d comme *decryption*.

L'étape astucieuse est de remarquer que

$$M^{1+k(p-1)(q-1)} \in [M]_p \cap [M]_q. \tag{20.29}$$

Pour montrer cela nous utilisons le petit théorème de Fermat 6.13 et la remarque 6.14.

— 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 (20.29). Le nombre $M^{1+k\varphi(n)} - M$ est donc à la fois multiple de p et de q.

Le lemme chinois 6.24 nous dit immédiatement [8] qu'alors

$$M^{1+k\varphi(n)} - M \tag{20.30}$$

est un multiple de $pq = n$, c'est-à-dire que

$$C^d = M^{ed} \in [M]_n. \tag{20.31}$$

Si on ne croit pas au lemme chinois, on peut utiliser le lemme de Gauss. Posons

$$M^{1+k\varphi(n)} - M = ap = bq. \tag{20.32}$$

Dans ce cas p divise bq, mais q est premier avec p, donc le lemme de Gauss 3.112 nous enseigne [9] que p divise b.

20.2.4 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 : vu que M est entre 0 et $n = pq$, les multiples de p possibles sont $p, 2p, \cdots pq$. Il y a dont 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.

Problèmes et choses à faire

Est-ce que ces trois points sont corrects ?

20.2.5 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.

8. 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.
9. Ici aussi, si $p = q$, ça ne marche pas.

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 [10].

Trouver des nombres u et v tels que $Au + Bv = \text{pgcd}(A, B)$ est un problème expliqué en 3.3.4.

En ce qui concerne le calcul de M^e lorsque e est grand, il n'est évidemment pas pensable de faire $M \cdot M \cdot \ldots M$ avec e facteurs. Un truc pour calculer en moins d'étapes est l'**exponentiation rapide**. Si $e = 2k$ est pair, nous calculons

$$M^e = (M^k)^2; \tag{20.33}$$

si $e = 2k + 1$ alors nous calculons

$$M^e = M(M^k)^2. \tag{20.34}$$

Le calcul prend alors seulement environ $\log_2(e)$ étapes. Pour donner une idée,

$$\log_2(10^{120}) \simeq 400. \tag{20.35}$$

Très raisonnable, mais un ordinateur reste indispensable.

20.2.6 La solidité de RSA

La solidité de la méthode repose sur deux conjectures (non démontrées !!) :

— Pour déchiffrer il *faut* connaitre p et q.

— La difficulté de trouver p et q en partant de $n = pq$ est exponentielle en n.

Dans la méthode de déchiffrage 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.

20.2.7 Note non mathématique pour doucher l'enthousiasme

Il est souvent dit[333] 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 [11]. La réalité est qu'il est souvent possible de craquer un système de chiffrement arbitrairement complexe, même sans connaitre le petit théorème de Fermat ...
... tout dépends du contexte.

20.3 Polynômes cyclotomiques

20.3.1 Définitions et propriétés

Définition 20.17.
*Le **polynôme cyclotomique** d'indice n est le polynôme*

$$\phi_n(X) = \prod_{z \in \Delta_n} (X - z) \tag{20.36}$$

où Δ_n est l'ensemble des racines primitives de l'unité de la définition 20.8 :

$$\Delta_n = \{e^{2ik\pi/n} \text{ tel que } 0 \leqslant k \leqslant n-1 \text{ tel que } \text{pgcd}(k, n) = 1\}, \tag{20.37}$$

10. Cela donne une idée des connaissances en math des klingons, dont le docteur Spock parvient à craquer le code mentalement en deux heures.

11. Cela dit, le navigateur Tor[334], qui est un pur produit de RSA, permet effectivement d'accéder en France aux sites bloqués pour apologie du terrorisme (mars 2015).

Le polynôme ϕ_n est un polynôme unitaire de degré $\varphi(n)$ où φ est l'indicatrice d'Euler[12] $\varphi(n)$. Nous avons par exemple

$$\Delta_1 = \{1\} \tag{20.38a}$$
$$\Delta_2 = \{-1\} \tag{20.38b}$$
$$\Delta_3 = \{e^{2\pi i/3}, e^{4\pi i/3}\} \tag{20.38c}$$

et les premiers polynômes cyclotomiques sont donnés par

$$\phi_1(X) = X - 1 \tag{20.39a}$$
$$\phi_2(X) = X + 1 \tag{20.39b}$$
$$\phi_3(X) = X^2 + X + 1. \tag{20.39c}$$

Pour le dernier nous avons utilisé le fait que $e^{6\pi i/3} = 1$ et $e^{4\pi i/3 + e^{2\pi i/3}} = -1$.

Lemme 20.18.
Le polynôme $X^n - 1$ se factorise des diverses manières suivantes :

$$X^n - 1 = \prod_{z\in \mathrm{U}_n} (X - z) = \prod_{d\,|\,n} \prod_{z\in\Delta_d} (X - z) = \prod_{d\,|\,n} \phi_d(X) \tag{20.40}$$

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 3.177 dit qu'il existe un nombre k (polynôme de degré zéro) tel que $X^n - 1 = k\prod_{z\in\mathrm{U}_n}(X - z)$. Vu le coefficient du terme de plus haut degré, ce k ne peut être que 1.

Pour la suite nous utilisons l'union disjointe $\mathrm{U}_n = \bigcup_{d\,|\,n} \Delta_d$ du lemme 20.10 et la définition (20.36) des polynômes cyclotomiques. □

Remarque 20.19.
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 $\mathrm{Card}(\Delta_d) = \varphi(d)$ et $\sum_{d\,|\,n}\varphi(d) = n$ (définition 20.8 et lemme 20.10).

Proposition 20.20.
Les polynômes cyclotomiques sont à coefficients entiers : $\phi_n \in \mathbb{Z}[X]$.

Démonstration. Nous devons démontrer que les coefficients de ϕ_n sont dans $\mathbb{Z}$ alors qu'ils sont a priori dans $\mathbb{C}$. Nous démontrons cela par récurrence. D'abord $\phi_1(X) = X - 1$, d'accord. Ensuite

$$X^{n+1} - 1 = \prod_{d\,|\,n+1} \phi_d(X) = \phi_{n+1}(X) \cdot \underbrace{\prod_{\substack{d\,|\,n+1\\ d\leqslant n}} \phi_d(X)}_{\in\mathbb{Z}[X] \text{ par récurrence}} \tag{20.41}$$

Le lemme 6.42 conclut que $\phi_{n+1} \in \mathbb{Z}[X]$. Nous avons vu $\mathbb{Z}$ comme sous anneau du corps $\mathbb{C}$. □

Proposition 20.21.
Soient $1 \leqslant m \leqslant n$ deux entiers et

$$T(X) = \frac{X^n - 1}{X^m - 1} \in \mathbb{Z}(X). \tag{20.42}$$

Alors :

(1) si $m \mid n$ alors $T \in \mathbb{Z}[X]$,

(2) si $m \mid n$ et si $m < n$ alors ϕ_n divise T dans $\mathbb{Z}[X]$.

12. Définie par l'équation 20.11.

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

$$Q = \{\text{diviseurs de } n \text{ ne divisant pas } m\}. \tag{20.43}$$

Nous avons alors

$$X^n - 1 = \prod_{d|n} \phi_d(X) = \prod_{d|m} \phi_d(X) \cdot \prod_{q\in Q} \phi_q(X) = (X^m - 1) \cdot \prod_{q\in Q} \phi_q(X). \tag{20.44}$$

Nous avons donc

$$T(X) = \frac{X^n - 1}{X^m - 1} = \prod_{q\in Q} \phi_q(X) \in \mathbb{Z}[X]. \tag{20.45}$$

(2) Nous venons de montrer que

$$T = \prod_{q\in Q} \phi_q \in \mathbb{Z}[X]. \tag{20.46}$$

Étant donné que $m < n$ nous avons $n \in Q$ et donc

$$T = \phi_n \cdot \prod_{q\in Q\setminus\{n\}} \phi_q. \tag{20.47}$$

Par conséquent ϕ_n divise T dans $\mathbb{Z}[X]$.

□

Corollaire 20.22.
Si p est premier alors le polynôme cyclotomique ϕ_p a une bonne tête :

$$\phi_p(X) = 1 + X + \cdots + X^{p-1}. \tag{20.48}$$

Démonstration. Nous utilisons la formule du lemme 20.18 en remarquant que seuls p et 1 divisent p :

$$X^p - 1 = \prod_{d|p} \phi_d(X) = \phi_1(X)\phi_p(X) = (X - 1)\phi_p(X). \tag{20.49}$$

Nous pouvons simplifier par $X - 1$ en utilisant la formule du lemme 3.181(2) :

$$1 + X + \cdots + X^{p-1} = \phi_p(X) \tag{20.50}$$

□

Proposition 20.23 (Irréductibilité des polynômes cyclotomiques[335])**.**
Les polynômes cyclotomiques sont irréductibles sur $\mathbb{Q}$.

Démonstration. Pour rappel, nous savons déjà que pour tout $n \in \mathbb{N}$, $\phi_n \in \mathbb{Z}[X]$. Vu que les racines de ϕ_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 ϕ_n) ; en effet vu que ces polynômes divisent ϕ_n, s'ils sont distincts, la proposition 6.96 s'applique et le produit des polynômes minimaux diviserait ϕ_n. Dans le cas inverse, ϕ_n est polynôme minimal des racines primitives de l'unité et est donc irréductible. Soit donc ξ, une telle racine primitive. Une autre racine primitive est de la forme ξ^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 ξ et ξ^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 ϕ_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 ϕ_n est irréductible dans $\mathbb{Q}[X]$. Quoi

qu'il en soit, le lemme de Gauss 6.44 nous montre que $h \in \mathbb{Z}[X]$ parce que ϕ_n, f et g ont des coefficients entiers. Nous avons

$$f(\xi) = g(\xi^l) = 0. \tag{20.51}$$

Considérons le polynôme $\psi(X) = g(X^l)$. Ce polynôme ψ est dans $\mathbb{Z}[X]$ et ψ est annulateur de ξ, donc f divise ψ en tant que polynôme minimal de ξ. Il y a un polynôme unitaire à coefficients entiers (lemme de Gauss forever) k tel que

$$\psi = fk \tag{20.52}$$

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

$$\bar{\psi}(X) = \bar{g}(X^l) = \bar{g}(X)^l. \tag{20.53}$$

Par conséquent dire que $\bar{f}$ divise $\bar{\psi}$ revient à dire que $\bar{f}(X)$ divise $\bar{g}(X)^l$. En particulier tous 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 ϕ_n. Dans un corps de décomposition de ce facteur, ϕ_n aurait une racine double, alors que ce n'est pas le cas. Contradiction. Nous concluons que $f = g$. □

Le corollaire suivant va être utilisé pour déterminer les polygones constructibles à la règle et au compas, théorème de Gauss-Wantzel 20.87.

Corollaire 20.24.
Soit p un nombre premier et α 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 ϕ_q.

Démonstration. Le polynôme ϕ_q est irréductible par la proposition 20.23, il est unitaire par définition et contient le monôme $X - e^{2i\pi/q}$, donc il est annulateur. Annulateur, irréductible et unitaire, la proposition 6.65(2) en fait le polynôme minimal de ω. □

Théorème 20.25.
Soit $P \in \mathbb{Z}[X]$ un polynôme unitaire irréductible non constant tel que toutes les racines dans $\mathbb{C}$ soient de module $\leqslant 1$. Alors $P = X$ ou P est un polynôme cyclotomique.

Démonstration. Nous supposons que $X \neq 0$, et 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

$$P = \prod_{i=1}^{d}(X - \xi_i) \tag{20.54}$$

avec $\prod_{i=1}^d \xi_i = a_0$. Par hypothèse, $|\xi_i| \leqslant 1$ et donc $0 < |a_0| \leqslant 1$. Vu que $P \in \mathbb{Z}[X]$ nous avons donc $a_0 = 1$ et donc $|\xi_i| = 1$ pour tout i.

Nous introduisons les polynômes

$$g_q(X) = \prod_{i=1}^{d}\big(X - (\xi_i)^q\big), \tag{20.55}$$

et en particulier $g_1 = P$, et nous développons

$$g_q(X) = X^n + C_{1,q}X^{n-1} + \cdots + C_{n,q} \tag{20.56}$$

où

$$C_{k,q} = (-1)^k \sum_{1\leqslant i_1<\ldots<i_k\leqslant d} (\xi_{i_1}\ldots\xi_{i_k})^q. \tag{20.57}$$

Nous introduisons aussi les polynômes

$$F_{k,q}(X_1,\ldots,X_n) = (-1)^k \sum_{1\leqslant i_1<\ldots<i_k\leqslant d} (X_{i_1}\ldots X_{i_k})^q \tag{20.58}$$

qui sont des polynômes symétriques. Ils vérifient deux propriétés. La première est que

$$C_{r,q} = F_{r,q}(\xi_1,\ldots,\xi_n), \tag{20.59}$$

et la seconde est que les polynômes $F_{r,1}$ sont les polynômes symétriques élémentaires à un coefficient près. Le théorème 6.153 nous donne alors des polynômes $G_{k,q}\in\mathbb{Z}[X_1,\ldots,X_n]$ tels que

$$F_{k,q}(X_1,\ldots,X_n) = G_{k,q}\big(F_{1,1}(X_1,\ldots,X_n),\ldots,F_{k,1}(X_1,\ldots,X_n)\big). \tag{20.60}$$

Nous savons que

$$|C_{k,q}| \leqslant \sum_{1\leqslant i_1<\ldots<i_k<d} 1 = \binom{d}{k}. \tag{20.61}$$

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

$$\max_{k=1,\ldots,d}\binom{d}{k}. \tag{20.62}$$

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, vu que $C_{1,q} = \sum_i \xi_i^q$, pour chaque k, l'ensemble

$$\{\xi_k^q \text{ tel que } q\in\mathbb{N}\}. \tag{20.63}$$

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 $N = \text{ppcm}(N_1,\ldots,N_d)$, nous avons

$$\xi_k^N = 1 \tag{20.64}$$

pour tout k.

Mais P est irréductible dans $\mathbb{Z}[X]$; s'il a ±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 ±1 parmi ses racines, alors P n'a pas de racines dans $\mathbb{Q}$ parce que ±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

$$X^N - 1 = \prod_{d\,|\,N}\phi_d(X), \tag{20.65}$$

les polynômes cyclotomiques sont les seuls facteurs irréductibles de X^N-1. Donc P est un polynôme cyclotomique. □

20.3.2 Nombres premiers

Lemme 20.26 ([336]).
Soit $n\geqslant 1$. 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

$$B(X) = \prod_{\substack{d \mid n \\ d \neq n}} \phi_d(X), \tag{20.66}$$

et nous commençons par montrer que ϕ_n est premier avec B. Nous avons $X^n - 1 = B\phi_n$, donc B et ϕ_n n'ont pas de racines communes (même pas dans $\mathbb{C}$) parce que ce serait une racine double de $X^n - 1$. Notons que par définition 20.36, les polynômes cyclotomiques sont scindés (dans $\mathbb{C}$), donc en particulier les polynômes ϕ_n et B sont scindés et donc premiers entre eux, dans $\mathbb{C}$ et a fortiori dans $\mathbb{Q}$. Par Bézout (corollaire 3.111), il existe $U, V \in \mathbb{Q}[X]$ tels que

$$U\phi_n + VB = 1. \tag{20.67}$$

Si nous prenons $a \in \mathbb{Z}$ tel que $U' = aU$ et $V' = aV$ soient tous deux dans $\mathbb{Z}[X]$, alors nous avons

$$U'\phi_n + V'B = a, \tag{20.68}$$

égalité dans $\mathbb{Z}[X]$. Quitte à prendre un multiple assez grand de a, nous pouvons choisir a de telle sorte que $|\phi_n(a)| \geqslant 2$. Nous prenons alors un nombre premier p divisant $\phi_n(a)$.

Montrons que le a et le p ainsi construits satisfont aux exigences.

Vu que $X^n - 1 = B\phi_n$, 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 (20.68) en a :

$$U'(a)\phi_n(a) + V'(a)B(a) = a. \tag{20.69}$$

Le nombre p ne divisant pas a, mais divisant $\phi_n(a)$, il ne peux pas diviser $B(a)$ [13]. 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 $\phi_n(a)$ entraine le fait de diviser $a^n - 1$ parce que ϕ_n est un des facteurs de $X^n - 1$. Soit maintenant $d \neq n$ divisant n ; nous avons

$$X^d - 1 = \prod_{d' \mid d} \phi_{d'}, \tag{20.70}$$

et cela est une partie du produit

$$\prod_{\substack{d \mid n \\ d \neq n}} \phi_d. \tag{20.71}$$

Vu que p ne divise aucun des $\phi_d(a)$ de ce dernier produit, a fortiori, il ne divise pas le produit 20.70, et donc pas $a^d - 1$. □

Lemme 20.27.
Si $n \geqslant 1$, alors il existe un nombre premier dans $[1]_n$, c'est-à-dire un nombre premier de la forme $1 + kn$ avec $k \in \mathbb{N}^*$.

Démonstration. Soient $n \geqslant 1$ et p, a les nombres donnés par le lemme 20.26. Vu que p divise $\phi_n(a)$, p divise $a^n - 1$ et donc $[a]_p$ a un ordre qui divise n dans $(\mathbb{Z}/p\mathbb{Z})^*$ parce que $[a]_p^n = [1]_p$.

Prenons $d \neq n$ divisant n. Nous savons que

$$a^d - 1 = \prod_{d' \mid d} \phi_{d'}(a). \tag{20.72}$$

13. C'est pour pouvoir dire ça que l'on a choisi $V' \in \mathbb{Z}[X]$ de telle sorte que $V'(a)$ soit dans $\mathbb{Z}$

Par construction de a et p, nous avons

$$[\phi_{d'}(a)]_p \neq 0 \tag{20.73}$$

Vu que $\mathbb{Z}/p\mathbb{Z}$ est intègre, le produit est également non nul, c'est-à-dire

$$\Big[\prod_{d'|d}\phi_{d'}(a)\Big]_p \neq 0, \tag{20.74}$$

et donc $[a]_p^a \neq 1$. Nous avons donc montré que si $d \neq n$ divise n, alors nous avons en même temps

$$[a]_p^n = 1 \tag{20.75}$$

et

$$[a]_p^d \neq 1. \tag{20.76}$$

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. □

Théorème 20.28 (Forme faible du théorème de Dirichlet [57]).
Pour tout $n \geqslant 1$, il existe une infinité de nombres premiers dans $[1]_n$.

Démonstration. Le lemme 20.27 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

$$N = np_1 \dots p_r. \tag{20.77}$$

Nous utilisons maintenant le lemme 20.27 avec ce N, c'est-à-dire qu'on a un nombre premier de la forme

$$p = 1 + kN = 1 + knp_1 \dots p_r. \tag{20.78}$$

Cela 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 $p_1, \dots, p_r$. □

20.4 Dénombrement des solutions d'une équation diophantienne

Théorème 20.29 ([57, 234]).
Soient des entiers naturels premiers dans leur ensemble[14] *$\alpha_1, \dots, \alpha_p$ et l'équation*

$$\alpha_1 n_1 + \cdots + \alpha_p n_p = n \tag{20.79}$$

pour les naturels n_i où n est un naturel donné. Nous notons S_n le nombre de solutions de cette équation. Alors :

(1) Il existe un algorithme (en temps fini) pour calculer S_n en fonction des α_i et de n.

(2) Nous avons le comportement asymptotique

$$S_n \sim \frac{1}{\alpha_1 \dots \alpha_p}\frac{n^{p-1}}{(p-1)!}. \tag{20.80}$$

Démonstration. Pour $|z| < 1$ dans $\mathbb{C}$, utilisant le lemme 16.28, nous écrivons le développement

$$F(z) = \prod_{i=1}^{p}\frac{1}{1-z^{\alpha_i}} = \prod_{i=1}^{p}\sum_{n\geqslant 0} z^{n\alpha_i}. \tag{20.81}$$

14. Définition 3.11.

Nous allons maintenant à la pêche au terme de degré k dans ce produit de sommes en utilisant p fois le produit de Cauchy de la formule (16.38). Nous avons

$$F(z) = \sum_{k\geqslant 0} \left(\sum_{n_1\alpha_1+\cdots+n_p\alpha_p=n} 1 \right) z^k = \sum_{k\geqslant 0} S_k z^k. \tag{20.82}$$

La technique pour déterminer la valeur de S_n est alors de développer $F(z)$ en série de façon un peu explicite et d'identifier le coefficient de z^n parce que nous venons de voir que ce coefficient est S_n. Nous commençons par une décomposition en éléments simples, expliquée autour de l'équation (20.22) :

$$\frac{1}{1-z^{\alpha_i}} = \sum_{\alpha\in U_{\alpha_i}} \frac{A_{\omega,i}}{\omega - z}. \tag{20.83}$$

où U_{α_i} est le groupe des racines α_i^e^de l'unité décrit en 20.1. La raison de ce développement est que, comme mentionné dans le lemme 20.18, $\prod_{\omega\in \mathrm{U}_{\alpha_i}}(z-\omega) = z^{\alpha_1}-1$. Lorsque nous effectuons la somme, le dénominateur commun est donc bien [15] $1-z^{\alpha_i}$. En récrivant le produit :

$$F(z) = \prod_{i=1}^{p} \frac{1}{1-z^{\alpha_i}} = \prod_{i=1}^{p} \sum_{\omega\in U_{\alpha_i}} \frac{A_{\omega,i}}{\omega - z} \tag{20.84}$$

Les coefficients $A_{\omega,i}$ sont calculables explicitement, en temps fini.

Vu que 1 est dans tous les U_{α_i}, le produit fait intervenir au dénominateur des puissances de $(1-z)$ jusqu'à la puissance p. Les autres racines de l'unité appartiennent au maximum à $p-1$ des groupes U_{α_i} parce que les nombres α_i sont premiers dans leur ensemble, voir la proposition 20.7.

La fonction F peut alors s'écrire sous la forme

$$F(z) = \frac{A}{(1-z)^p} + G(z) \tag{20.85}$$

où $G(z)$ est une somme de termes de la forme

$$\frac{a_{i,1}}{1-\omega_i} + \cdots + \frac{a_{i,p}}{(1-\omega_i)^{p-1}} \tag{20.86}$$

où les ω_i sont les racines α_i^e^de l'unité et $a_{k,r}$ sont des nombres complexes. Trouvons A. D'abord grâce au lemme 3.181(2) nous avons

$$F(z)(1-z)^p = \prod_{l=1}^{p} \frac{1-z}{1-z^{\alpha_i}} = \prod_{i=1}^{p} \frac{1}{1+z+\cdots+z^{\alpha_i-1}}, \tag{20.87}$$

et donc

$$\lim_{z\to 1} F(z)(1-z)^p = \prod_{i=1}^{p} \frac{1}{\alpha_i}. \tag{20.88}$$

Mais vu ce que contient $G(z)$, nous avons aussi $\lim_{z\to 1} F(z) = A$. Nous avons donc déjà déterminé $A = \frac{1}{\alpha_1\ldots\alpha_p}$.

Pour la suite nous avons besoin des développements du lemme 16.28. Nous utiliserons en particulier celle-ci :

$$\frac{1}{(\omega-z)^k} = \frac{1}{(k-1)!} \sum_{s=0}^{\infty} \omega^{-s-1-k} \frac{(s+k-1)!}{s!} z^s. \tag{20.89}$$

En particulier le module du coefficient de z^n là dedans est : $\frac{(n+k-1)!}{n!(k-1)!}$. Dans la partie G de la décomposition (20.85), k est majoré par $p-2$ et la dépendance en n est donc au maximum du type

$$\frac{(n+p-2)!}{n!(p-2)!} \sim \frac{n^{n+p-2}}{n^n(p-2)!} = \frac{n^{p-2}}{(p-2)!}. \tag{20.90}$$

15. Pour le signe, c'est ajustable avec le signe de $A_{\omega,i}$.

Dans le premier terme par contre, il y a des termes jusqu'à $k = p$. Le terme dominant est alors en $\frac{n^{p-1}}{(p-1)!}$ et son coefficient est A qui est déjà calculé. Au final le terme dominant du coefficient de z^n dans $F(z)$ est

$$S_n \sim \frac{A}{(p-1)!}n^{p-1} = \frac{1}{\alpha_1 \dots \alpha_p}\frac{n^{p-1}}{(p-1)!}. \tag{20.91}$$

□

Exemple 20.30
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 α divise n, c'est-à-dire que n soit un multiple de α. Il n'y a que un nombre sur α à être multiple de α. D'où le comportement en $\frac{1}{\alpha}$.

Pour $p = 2$, c'est l'équation (3.58) déjà étudiée. Il y a une famille à un paramètre de solutions dont seulement un certain nombre sont positives. A priori, le nombre de solutions positives croît linéairement en n. △

20.5 Corps finis

Si vous cherchez des choses à propos de RSA, c'est à la section 20.2.

20.5.1 Théorème de Wedderburn

Théorème 20.31 (Théorème de Wedderburn[85])**.**
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 6.50 nous avons

$$\text{Card}(\mathbb{K}) = q^n \tag{20.92}$$

pour un certain n.

Nous supposons maintenant que $\mathbb{K}$ est non commutatif. Dans ce cas $Z \neq \mathbb{K}$ et nous avons $n \geqslant 2$. Nous considérons aussi

$$Z_x = \{a \in \mathbb{K} \text{ tel que } ax = xa\}. \tag{20.93}$$

Le centre Z est un sous corps de Z_x, donc il existe $d(x)$ tel que

$$\text{Card}(Z_x) = q^{d(x)}. \tag{20.94}$$

De la même manière, Z_x est un sous corps de $\mathbb{K}$, donc il existe $m(x)$ tel que

$$\text{Card}(\mathbb{K}) = \text{Card}(Z_x)^{m(x)}. \tag{20.95}$$

En mettant bout à bout nous avons

$$q^n = \text{Card}(Z_x)^{m(x)} = q^{d(x)m(x)}, \tag{20.96}$$

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 $\mathbb{K}^*$ sur lui-même :

$$\varphi(k)x = kxk^{-1}. \tag{20.97}$$

Nous notons $\mathcal{O}_x$ l'orbite de $x \in \mathbb{K}^*$ pour cette action, et $\text{Fix}(x)$ son stabilisateur. Nous avons

$$Z_y = \text{Fix}(y) \cup \{0\} \tag{20.98}$$

parce que Z_y et $\mathrm{Fix}(y)$ ont les mêmes définitions, sauf que $\mathrm{Fix}(y)$ est dans $\mathbb{K}^*$ alors que Z_y est dans $\mathbb{K}$. Nous avons donc

$$\mathrm{Card}\big(\mathrm{Fix}(y)\big) = \mathrm{Card}(Z_y) - 1 = q^{d(y)} - 1. \tag{20.99}$$

Nous avons $\mathrm{Card}(\mathcal{O}_x) = 1$ si et seulement si $\mathcal{O}_x = \{x\}$ si et seulement si $\mathrm{Fix}(x) = \mathbb{K}^*$ si et seulement si $z \in Z^*$. 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^* sont

$$\{z_1\}, \dots, \{z_{q-1}\} \tag{20.100}$$

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 (2.88) :

$$\mathrm{Card}(\mathbb{K}^*) = \mathrm{Card}(Z^*) + \sum_{i=1}^{r} \frac{\mathrm{Card}(\mathbb{K}^*)}{\mathrm{Card}(\mathrm{Fix}(y_i))}, \tag{20.101}$$

mais $\mathrm{Card}(Z^*) = q - 1$, $\mathrm{Card}(\mathbb{K}^*) = q^n - 1$ et $\mathrm{Card}\big(\mathrm{Fix}(y_i)\big) = q^{d(y_i)} - 1$, donc

$$q^n - 1 = (q - 1) + \sum_{i=1}^{r} \frac{q^n - 1}{q^{d(y_i)} - 1}. \tag{20.102}$$

Nous considérons la fraction rationnelle

$$F(X) = (X^n - 1) - \sum_{i=1}^{r} \frac{X^n - 1}{X^{d(y_i)} - 1}. \tag{20.103}$$

Étant donné que $d(y_i)$ divise n, nous avons, contrairement aux apparences, que $F \in \mathbb{Z}[X]$ par la proposition 20.21(1).

Nous pouvons exploiter un peu mieux la proposition 20.21 en remarquant que $d(y_i) < n$ parce que sinon $\mathrm{Card}(Z_{y_i}) = \mathrm{Card}(\mathbb{K})$, ce qui signifierait que $y_i \in Z$, ce qui nous avions exclu. Par conséquent le polynôme cyclotomique ϕ_n divise

$$\frac{X^n - 1}{X^{d(y_i)} - 1} \tag{20.104}$$

dans $\mathbb{Z}[X]$. Le polynôme cyclotomique ϕ_n divise également $X^n - 1$ et par conséquent ϕ_n divise F. Il existe donc $Q \in \mathbb{Z}[X]$ tel que $F = Q\phi_n$. En particulier en évaluant en q :

$$F(q) = Q(q)\phi_n(q) = q - 1. \tag{20.105}$$

En effet nous avons $F(q) = q - 1$ par construction : comparer (20.102) avec (20.103). Évidemment $q \neq 1$ parce que si $q = 1$ alors $\mathrm{Card}(\mathbb{K}) = 1$ 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 (20.105) nous avons $q - 1 \neq 0$. Nous avons donc $|Q(q)| \geqslant 1$ et donc

$$|\phi_n(q)| \leqslant q - 1. \tag{20.106}$$

Par définition du polynôme cyclotomique nous avons

$$|\phi_n(q)| = \prod_{z \in \Delta_n} |q - z|. \tag{20.107}$$

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 $|z_0 - q| \leqslant q - 1$. Étant donné que $n \geqslant 2$ nous avons $z_0 \neq 1$.

Mais d'autre part, comme indiqué sur la figure 20.1, 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 $z = 1$.

Nous avons ainsi obtenu une contradiction, et nous concluons que le corps $\mathbb{K}$ est commutatif. □

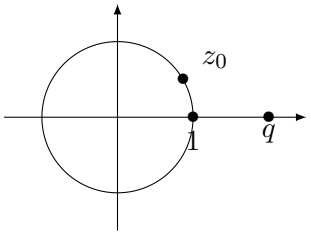


FIGURE 20.1 – Nous devons avoir $|z_0 - q| > q - 1$.

20.5.2 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.

Théorème 20.32.
Soit p un nombre premier, soit $n \in \mathbb{N}^*$ 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 corollaire 2.32 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

$$X^{q-1} - 1 \tag{20.108}$$

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 6.122.

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

$$(\alpha\beta)^q = \alpha^q \beta^q = \alpha\beta \tag{20.109}$$

parce que $\alpha^q = \alpha$. Le produit $\alpha\beta$ est donc encore dans $\mathbb{E}$. Pour la somme,

$$(\alpha+\beta)^q = (\alpha+\beta)^{p^n} = \Big((\alpha+\beta)^p\Big)^{p^{n-1}} = (\alpha^p + \beta^p)^{p^{n-1}} = \ldots = \alpha^{p^n} + \beta^{p^n} = \alpha + \beta. \tag{20.110}$$

En ce qui concerne l'inverse,

$$(\alpha^{-1})^q = (\alpha^q)^{-1} = \alpha^{-1}. \tag{20.111}$$

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 α_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

$$P'(X) = qX^{q-1} - 1 = -1 \tag{20.112}$$

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 racines doubles. Toutes les racines étant simples, il y en a exactement q.

□

Le théorème 20.32 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 construire. Ces résultats proviennent de [337, 338, 339] et de wikipedia

Proposition 20.33 ([339]).
Soit $\mathbb{K}$ un corps fini. Alors le groupe multiplicatif $\mathbb{K}^*$ est cyclique.

Démonstration. Soit $\mathbb{K}$ un corps ayant q éléments. Le groupe $\mathbb{K}^*$ en a $q-1$; ergo l'ordre des éléments de $\mathbb{K}^*$ sont des diviseurs de $q-1$; c'est le corollaire 2.32. Soit d un diviseur de $q-1$ et

$$H_d^* = \{x \text{ d'ordre } d \text{ dans } \mathbb{K}^*\} \tag{20.113a}$$
$$H_d = \{\text{racines de } X^d - 1 \text{ dans } \mathbb{K}\}. \tag{20.113b}$$

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 $H_d^* \neq \varnothing$ et considérons $a \in H_d^*$. Alors l'application

$$\begin{aligned} \phi\colon \mathbb{Z}/d\mathbb{Z} &\to H_d \\ n &\mapsto a^n \end{aligned} \tag{20.114}$$

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 $\text{Card}(H_d) \leqslant d$ 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 ϕ 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 ϕ.

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 20.12, et H_d étant isomorphe à $\mathbb{Z}/d\mathbb{Z}$ nous avons

$$\text{Card}(H_d^*) = \varphi(d). \tag{20.115}$$

Par conséquent si H_d^* n'est pas vide, son cardinal est $\varphi(d)$. Nous avons

$$\begin{aligned} q - 1 &= \text{Card}(\mathbb{K}^*) && \text{(20.116a)} \\ &= \text{Card}\Big(\bigcup_{d|q-1} H_d^*\Big) && \text{(20.116b)} \\ &= \sum_{d|q-1} \text{Card}(H_d^*) && \text{(20.116c)} \\ &\leqslant \sum_{d|q-1} \varphi(d) && \text{(20.116d)} \\ &= q - 1 && \text{(20.116e)} \end{aligned}$$

où nous avons utilisé le lemme 20.10. 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. □

Corollaire 20.34.
Si p est un nombre premier, alors

$$(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}. \tag{20.117}$$

L'isomorphisme est un isomorphisme de groupes (abéliens). À gauche multiplicatif et à droite additif.

Démonstration. La proposition 20.33 nous enseigne que le 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 $n = p - 1$. □

Lorsque $\mathbb{K}$ est un corps les éléments du groupe $\mathbb{K}^*$ sont les **éléments primitifs** de $\mathbb{K}$.

Proposition 20.35.
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 $\mathbb{K}^*$ ayant $q-1$ éléments, ses éléments vérifient $a^{q-1} = 1$ par le corollaire 2.32 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

$$\prod_{a\in\mathbb{K}}(X - a) \tag{20.118}$$

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. □

Exemple 20.36
Soit $\mathbb{K} = \mathbb{Q}$ et $\mathbb{L} = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. Afin de montrer que $\mathbb{L} = \mathbb{Q}(\alpha)$ avec $\alpha = \sqrt{2} + \sqrt{3}$ nous devons montrer que $\sqrt{2}$ et $\sqrt{3}$ sont des polynômes en α. △

Une conséquence du fait que $x^q = x$ est qu'il ne faut pas regarder le théorème 6.99 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.

Exemple 20.37
Si $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$ est le[16] corps à 3 éléments, alors le polynôme $P(X) = X^3 - X$ s'évalue à zéro pour tout $x \in \mathbb{F}_3$ (proposition 20.35.) mais il n'est pas le polynôme nul. △

20.5.3 Symboles de Legendre et carrés

Source : [340].
Nous disons que $a \in \mathbb{F}_p$ est un **carré** s'il existe $b \in \mathbb{F}_p$ tel que $a = b^2$.

Définition 20.38.
*Soit $n \in \mathbb{N}$ et $p > 2$ un nombre premier. Le **symbole de Legendre** par*

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & \text{si } p \text{ divise } n \\ 1 & \text{si } n \text{ est un carré dans } \mathbb{F}_p \\ -1 & \text{sinon.} \end{cases} \tag{20.119}$$

Note que -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.

Proposition 20.39.
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

$$\left(\frac{n}{p}\right) = n^{(p-1)/2} \mod p. \tag{20.120}$$

Démonstration. Nous considérons l'application

$$\begin{aligned} \psi\colon \mathbb{F}_p^* &\to \mathbb{F}_p^* \\ x &\mapsto x^2. \end{aligned} \tag{20.121}$$

16. Le singulier est justifié par le théorème 20.32, mais ça n'a pas d'importance ici.

C'est un morphisme de groupes multiplicatifs et $\ker\psi = \{-1,1\}$. Étant donné que $p>2$, nous avons alors

$$\text{Card}(\ker\psi) = 2 \tag{20.122}$$

parce que $1\neq -1$. Évidemment l'ensemble des carrés dans $\mathbb{F}_p^*$ est l'image de ψ. Le premier théorème d'isomorphisme 2.25(3) nous permet alors de conclure que

$$\text{Card}(\text{Image}(\psi)) = \frac{\text{Card}(\mathbb{F}_p^*)}{2}. \tag{20.123}$$

Ceci prouve la première assertion.

Par le petit théorème de Fermat (théorème 6.13), 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

$$X^{(p-1)/2} = \pm 1. \tag{20.124}$$

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 (20.120). 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

$$n^{(p-1)/2} = n^{p-1} = 1 = \left(\frac{n}{p}\right). \tag{20.125}$$

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

$$n^{(p-1)/2} = -1 = \left(\frac{n}{p}\right). \tag{20.126}$$

□

Corollaire 20.40.
Si $a,b\in\mathbb{N}$ et si $p>2$ est un nombre premier, alors

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{20.127}$$

Démonstration. Par la formule (20.120),

$$\left(\frac{ab}{p}\right) = (ab)^{(p-1)/2} = a^{(p-1)/2}b^{(p-1)/2} = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{20.128}$$

□

Soit un nombre premier $q>2$ et $\mathbb{A}$, un anneau de caractéristique p. Si $\alpha\in\mathbb{A}$ vérifie

$$1+\alpha+\cdots+\alpha^{q-1}=0, \tag{20.129}$$

nous définissons la **somme de Gauss** par

$$\tau = \sum_{x\in\mathbb{F}_q}\left(\frac{i}{q}\right)\alpha^i = \sum_{x=1}^{q-1}\left(\frac{x}{q}\right)\alpha^i. \tag{20.130}$$

Notons que la somme de Gauss dépend de q et du α choisis.

Proposition 20.41.
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 $p \geqslant 3$ et si $p \neq q$ alors

$$\tau^p = \left(\frac{p}{q}\right)\tau. \tag{20.131}$$

(3) Si $\mathbb{A}$ est de caractéristique p et si q est premier avec p, alors τ est inversible dans $\mathbb{A}$.

Démonstration. D'abord nous notons que

$$\alpha^q - 1 = (\alpha - 1)(1 + \alpha + \cdots + \alpha^{q-1}) = 0 \tag{20.132}$$

par définition de α. Nous calculons

$$\epsilon(q)\tau^2 = \epsilon(q) \sum_{x,y\in\mathbb{F}_q} \left(\frac{x}{q}\right)\left(\frac{y}{q}\right)\alpha^{x+y} \tag{20.133a}$$

$$= \sum_{x,y\in\mathbb{F}_q} \left(\frac{-xy}{q}\right)\alpha^{x+y}. \tag{20.133b}$$

$$= \sum_{z\in\mathbb{F}_q}\sum_{y\in\mathbb{F}_q} \left(\frac{-(z-y)y}{q}\right)\alpha^{z} \tag{20.133c}$$

$$= \sum_{z\in\mathbb{F}_q} s_z\alpha^z \tag{20.133d}$$

Justifications :

— Pour obtenir (20.133b) nous avons utilisé le corollaire 20.40.

— (20.133c) est un changement de variable $z = x + y$ dans la somme sur x.

— Pour (20.133d) nous avons posé

$$s_z = \sum_{y\in\mathbb{F}_q} \left(\frac{-(z-y)y}{q}\right). \tag{20.134}$$

Nous avons

$$s_0 = \sum_{y\in\mathbb{F}_q} \left(\frac{y^2}{q}\right). \tag{20.135}$$

Dans cette somme, tous les termes sont 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

$$\begin{aligned} \mathbb{F}_q^* &\to \mathbb{F}_q\backslash\{1\} \\ k &\mapsto 1 - zy^{-1} \end{aligned} \tag{20.136}$$

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

$$\sum_{y\in\mathbb{F}_q} \left(\frac{y(z-y)}{q}\right) = \sum_{y\in\mathbb{F}_q} \left(\frac{y^2(y^{-1}z-1)}{q}\right) \tag{20.137a}$$

$$= \sum_{y\in\mathbb{F}_q} \left(\frac{y^{-1}z-1}{q}\right) \tag{20.137b}$$

$$= \sum_{t\in\mathbb{F}_q\backslash\{1\}} \left(\frac{t}{q}\right) \tag{20.137c}$$

$$= \underbrace{\sum_{t\in\mathbb{F}_q} \left(\frac{y}{q}\right)}_{=0} - \left(\frac{1}{1}\right) \tag{20.137d}$$

$$= -1 \tag{20.137e}$$

parce qu'il y a autant de carrés que de non carrés dans $\mathbb{F}_q^*$ (proposition 20.39). En résumé nous avons

$$\epsilon(q)\tau^2 = \sum_{z\in\mathbb{F}_q} s_z\alpha^z \tag{20.138}$$

où

$$s_z = \begin{cases} q-1 & \text{si } z=0 \\ -1 & \text{sinon.} \end{cases} \tag{20.139}$$

Cela donne

$$\epsilon(q)\tau^2 = (q-1) - \underbrace{(\alpha+\cdots+\alpha^{q-1})}_{=-1} = q \tag{20.140}$$

où nous avons utilisé l'hypothèse sur α. Donc $\epsilon(q)\tau^2 = q$, et étant donné que $\epsilon(q) = \pm 1$ nous concluons

$$\tau^2 = \epsilon(q)q. \tag{20.141}$$

Nous prouvons maintenant la seconde partie. Vu que $\mathbb{A}$ est de caractéristique p en utilisant le fait que le morphisme de Frobenius est un morphisme,

$$\tau^p = \left(\sum_{x\in\mathbb{F}_q}\left(\frac{x}{q}\right)\alpha^x\right)^p = \sum_{x\in\mathbb{F}_q}\left(\frac{x}{q}\right)^p\alpha^{px}. \tag{20.142}$$

Étant donné que $\left(\frac{x}{q}\right) = \pm 1$ et que p est impair, nous avons

$$\left(\frac{x}{q}\right)^p = \left(\frac{x}{q}\right). \tag{20.143}$$

Du coup nous avons

$$\left(\frac{p}{q}\right)\tau^p = \sum_{x\in\mathbb{F}_p}\left(\frac{xp}{q}\right)\alpha^{px}. \tag{20.144}$$

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 :

$$\left(\frac{p}{q}\right)\tau^p = \sum_{x\in\mathbb{F}_p}\left(\frac{x}{q}\right)\alpha^x = \tau. \tag{20.145}$$

Nous trouvons alors que

$$\tau^p = \left(\frac{p}{q}\right)\tau. \tag{20.146}$$

Étant donné la formule du τ^2 que nous venons de démontrer, nous avons $\tau^2 = \pm q$. Les nombres p et q étant premiers entre eux, la relation de Bézout (théorème 3.13) nous donne a et b tels que

$$ap + ba = 1. \tag{20.147}$$

Cela montre que b est un inverse de q modulo p. Donc τ^2 est inversible, et il en découle que τ lui-même est inversible. □

Théorème 20.42 (Loi de réciprocité quadratique)**.**
Soient deux nombres premiers distincts $p, q \geqslant 3$. Alors

$$\left(\frac{p}{q}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}\left(\frac{q}{p}\right). \tag{20.148}$$

Démonstration. Soit ϕ_q le polynôme $1 + X + \cdots + X^{q-1}$ et l'anneau

$$\mathbb{A} = \mathbb{F}_p[X]/(\phi_q). \tag{20.149}$$

Cela 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

$$\tau = \sum_{i \in \mathbb{F}_q} \left(\frac{i}{q}\right) \alpha^i. \tag{20.150}$$

Notons que cela 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 polyôme. Donc les coefficients de α doivent être compris comme des éléments de $\mathbb{F}_p$. Nous savons (proposition 20.41) que

$$\tau^2 = \left(\frac{-1}{q}\right), \tag{20.151}$$

et en utilisant la formule (20.120) nous trouvons

$$\left(\frac{\tau^2}{p}\right) = (\tau^2)^{(p-1)/2} \mod p = \tau^{p-1} \mod p \tag{20.152}$$

En réalité sur cette dernière ligne, nous ne devrions pas préciser le « mod p » parce que, comme mentionné plus haut, ce sont des éléments de $\mathbb{F}_p$. En utilisant cela ainsi que (20.131) nous avons

$$\underbrace{\left(\frac{\tau^2}{p}\right)}_{\tau^{p-1}} \tau = \tau^p = \left(\frac{p}{q}\right) \tau \tag{20.153}$$

Vu que τ est inversible, nous écrivons

$$\left(\frac{\tau^2}{p}\right) = \left(\frac{p}{q}\right) \tag{20.154}$$

Nous utilisons maintenant la formule (20.120) sur le membre de gauche avec $n = \tau^2 = \left(\frac{-1}{q}\right)$:

$$\left(\frac{p}{q}\right) = \left(\frac{-1}{q}\right)^{\frac{q-1}{2}} \left(\frac{q}{p}\right). \tag{20.155}$$

Toujours avec la même formule nous pouvons substituer $\left(\frac{-1}{q}\right)$ par $(-1)^{(q-1)/2}$ et obtenir

$$\left(\frac{p}{q}\right) = (-1)^{\frac{(q-1)}{2}\frac{(p-1)}{2}}. \tag{20.156}$$

□

Lemme 20.43.
Si p est un nombre premier $p \geqslant 3$, 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 20.39. Pour l'unicité, soit $\alpha : \mathbb{F}_p^* \to \{-1, 1\}$ un morphisme surjectif (c'est-à-dire non trivial). Étant donné que

$$\mathbb{F}_p^* = \ker(\alpha) \cup -\ker(\alpha), \tag{20.157}$$

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 $\mathbb{F}_p^*$ ne possède qu'un seul sous-groupe d'indice 2. En effet soit S un tel sous-groupe et a, un générateur de $\mathbb{F}_p^*$ (qui est cyclique par la proposition 20.33), alors $a^2 \in S$ par le lemme 3.30. 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 $\mathbb{F}_p^*$ est pair.

Bref, le sous-groupe $\ker(\alpha)$ est l'unique sous-groupe d'indice 2 dans $\mathbb{F}_p^*$. Mais la proposition 20.39 nous indique que $|(\mathbb{F}_p^*)^2| = \frac{p-1}{2}$, c'est-à-dire que le groupe des carrés est d'indice 2. Nous avons donc, par l'unicité,

$$\ker(\alpha) = (\mathbb{F}_p^*)^2. \tag{20.158}$$

Au final, pour $y \in \mathbb{F}_p^*$,

$$\alpha(y) = \begin{cases} 1 & \text{si } y \text{ est un carré} \\ -1 & \text{sinon.} \end{cases} \tag{20.159}$$

Cela est bien la définition des symboles de Legendre. □

Proposition 20.44.
Pour p premier nous avons

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{si } p \in [1]_8 \text{ ou } p \in [7]_8 \\ -1 & \text{sinon.} \end{cases} \tag{20.160}$$

Démonstration. Soit le polynôme

$$X^4 + 1 \in \mathbb{F}_p[X] \tag{20.161}$$

et α, une racine dans une extension [17] de $\mathbb{F}_p$ [18]. Nous posons $\theta = \alpha + \alpha^{-1}$ et nous calculons

$$\theta^2 = (\alpha + \alpha^{-1})(\alpha + \alpha^{-1}) = \alpha^2 + 2 + (\alpha^2)^{-1} = \alpha^2 + 2 - \alpha^2 = 2 \tag{20.162}$$

parce que α^2 étant -1, nous avons $(\alpha^2)^{-1} = -\alpha^2$. Bref, $\theta^2 = 2$.

Dire que 2 est un carré modulo p revient à dire que θ 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 θ est vraiment dans $\mathbb{F}_p$ et non seulement dans l'extension $\mathbb{F}_p(\alpha)$. En tenant compte de l'exemple 6.94, 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 α est :

$$1, \alpha, \alpha^2, \alpha^3, -1, -\alpha, -\alpha^2, -\alpha^3, 1, \alpha, \ldots \tag{20.163}$$

Nous avons donc automatiquement $\alpha^{9k} = \alpha$, mais $p = 9k$ est exclu parce que p est premier. Nous devons donc vérifier si une des possibilités

$$\alpha^2 = \alpha \tag{20.164a}$$
$$\alpha^3 = \alpha \tag{20.164b}$$
$$-\alpha = \alpha \tag{20.164c}$$
$$-\alpha^3 = \alpha \tag{20.164d}$$

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 6.94, 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,

$$\theta^p = (\alpha + \alpha^{-1})^p = \alpha^p + \alpha^{-p}. \tag{20.165}$$

17. Dans la source que je suivais (je ne sais plus où), on parlait ici de « fermeture » de $\mathbb{F}_p j$ et non d'exetension. Il me semble que parler simplement d'extension suffit. Vous confirmez ?

18. Voir par exemple la proposition 6.119 pour l'existence d'une extension comme il faut.

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

$$\theta^p = \alpha^{1+8k} + (\alpha^{-1})^{1+8k} = \alpha + \alpha^{-1} = \theta, \tag{20.166}$$

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

$$\alpha^6 + 1 = \alpha^4 + \alpha^2, \tag{20.167}$$

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. □

20.5.4 Théorème de Chevalley-Warning

Lemme 20.45.
Soit $\mathbb{K}$ un corps de caractéristique p et de cardinal q. Pour $m \in \mathbb{N}$ nous définissons

$$S_m = \sum_{x \in \mathbb{K}} x^m. \tag{20.168}$$

Alors nous avons

$$S_m \mod p = \begin{cases} -1 & \text{si } m \geqslant 1 \text{ et } m \text{ divisible par } q-1 \\ 0 & \text{sinon.} \end{cases} \tag{20.169}$$

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 3.56).

Nous prenons maintenant $m \geqslant 1$ 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

$$x^m = x^{k(q-1)} = 1 \tag{20.170}$$

parce que $\mathbb{K}^*$ est cyclique et $x^{q-1} = 1$ par le petit théorème de Fermat (théorème 6.13). Par conséquent nous avons

$$\sum_{x \in \mathbb{K}} x^m = \sum_{x \in \mathbb{K}^*} 1 = q - 1. \tag{20.171}$$

Si le nombre $m \geqslant 1$ n'est pas divisible par $q - 1$ alors nous prenons un générateur y du groupe $\mathbb{K}^*$. Un tel élément vérifie $y^m \neq 1$. En effet, si y vérifiait $y^m = 1$ alors cela signifierait que l'ordre de $\mathbb{K}^*$ est un diviseur de m, ce qui n'est pas le cas ici parce que l'ordre de $\mathbb{K}^*$ est $q - 1$. Pour un tel y, l'application

$$\begin{aligned} \varphi\colon \mathbb{K}^* &\to \mathbb{K}^* \\ x &\mapsto yx \end{aligned} \tag{20.172}$$

est une bijection [19]. 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

$$z = \varphi(a^{l-k}). \tag{20.173}$$

Nous pouvons maintenant faire le calcul.

$$S_m = \sum_{x \in \mathbb{K}^*} x^n = \sum_{x \in \mathbb{K}^*} (yx)^m = y^m \sum_{x \in \mathbb{K}^*} x^m = y^m S_m. \tag{20.174}$$

Étant donné que $y^m \neq 1$, la seule solution est $S_m = 0$. □

19. 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 $y \neq 0$.

Théorème 20.46 (Chevalley-Warninig[341]).
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 :

$$V = \{x \in \mathbb{K}^n \text{ tel que } P_1(x) = \dots = P_r(x) = 0\}. \tag{20.175}$$

Alors $\text{Card}(V) = 0 \mod p$.

Démonstration. Nous considérons le polynôme

$$P = \prod_{i=1}^{r} (1 - P_i^{q-1}). \tag{20.176}$$

Montrons que

$$P(x) = \begin{cases} 1 & \text{si } x \in V \\ 0 & \text{sinon.} \end{cases} \tag{20.177}$$

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 $P_i(x) \in \mathbb{K}^*$. Mais dans ce cas (toujours la cyclicité de $\mathbb{K}^*$) 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 que

$$\deg(P) = \sum_{i=1}^{r} (q-1) \deg(P_i) < n(q-1). \tag{20.178}$$

Pour un polynôme $Q \in \mathbb{K}[X_1, \dots, X_n]$, nous définissons

$$\int Q = \sum_{x \in \mathbb{K}^n} Q(x). \tag{20.179}$$

Nous avons immédiatement

$$\int P = \sum_{x \in \mathbb{K}^n} P(x) = \sum_{x \in V} 1 = \text{Card}(V) \mod p. \tag{20.180}$$

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

$$P = \sum_m c_m X_1^{m_1} \dots X_n^{m_n} \tag{20.181}$$

où la somme s'étend sur les $m \in \mathbb{N}^n$ tels que $c_m \neq 0$. Nous avons

$$\int P = \sum_{\in \mathbb{K}^n} \sum_m c_m x_1^{m_1} \dots x_n^{m_n} \tag{20.182a}$$

$$= \sum_m c_m \left(\sum_{x \in \mathbb{K}^n} x_1^{m_1} \dots x_n^{m_n} \right) \tag{20.182b}$$

$$= \sum_m c_m S_{m_1} \dots S_{m_n}. \tag{20.182c}$$

Le terme de plus haut degré dans la décomposition (20.181) est celui du m tel que $\sum_i m_i$ est le plus grand, mais étant donné que nous savons que ce degré est plus petit que $n(q-1)$, nous avons pour tous les m entrant dans la somme que

$$\sum_{i=1}^{n} m_i < n(q-1). \tag{20.183}$$

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

$$\sum_{m\in\mathbb{N}^n} c_m S_{m_1} \dots S_{m_n} \tag{20.184}$$

ont un facteur nul. □

Corollaire 20.47.
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 termes constants, alors ils ont un zéro commun non trivial.

Démonstration. Nous reprenons les notations du théorème 20.46. Étant donné que les P_i n'ont pas de termes constants, $0 \in V$, mais $\text{Card}(V) = 0 \mod p$. Par conséquent nous devons avoir $\text{Card}(V) > p$. □

Exemple 20.48
Nous considérons les polynômes

$$P_1(x, y, t, u) = xy + x + ux \tag{20.185a}$$
$$P_2(x, y, t, u) = x + y - 3t. \tag{20.185b}$$

La somme de leurs degrés est 3 et ce sont des polynômes à 4 variables. Nous devons donc avoir, en vertu du corollaire 20.47, des autres racines que la racine triviale $(x, y, t, u) = (0, 0, 0, 0)$.

Le corollaire nous donne aussi une borne inférieure 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 $P_1 = P_2 = 0$. △

20.5.5 Contenu d'un polynôme

Lemme 20.49 (de Gauss[43, 342])**.**
Soient $P, Q \in \mathbb{Z}[X]$. Alors

$$c(PQ) = c(P)c(Q) \tag{20.186}$$

où c est le contenu, définition 3.166.

Démonstration. Afin de fixer les notations, nous posons $P = \sum_i a_i X^i$ et $Q = \sum_j b_j Y^j$.

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 $c(PQ)$. Vu que le contenu de P et de Q sont 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 :

$$p \mid a_0, p \mid a_1, \dots, p \mid a_{i_0-1}, p \nmid a_{i_0} \tag{20.187}$$

et 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 :

$$a_{i_0}b_{j_0} + \sum_{\substack{i+j=i_0+j_0\\ i<i_0 \text{ ou } j<j_0}} a_i b_j. \tag{20.188}$$

Par définition p divise soit a_i soit b_j pour chacun des termes de la grande somme. Vu que 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 $c(PQ) = 1$.

Cas général Si P et Q sont maintenant des polynômes sans conditions particulières 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 :

$$c(P_1Q_1) = 1. \tag{20.189}$$

Étant donné que

$$P_1Q_1 = \frac{1}{c(P)c(Q)}PQ, \tag{20.190}$$

nous avons

$$c(PQ) = c(P)c(Q)c(P_1Q_1) = c(P)c(Q). \tag{20.191}$$

□

20.5.6 Théorème de l'élément primitif

Définition 20.50.
*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.

Théorème 20.51 (de l'élément primitif)**.**
Si $\mathbb{K}$ est un corps fini, toute extension finie de $\mathbb{K}$ est simple[20].
Si $\mathbb{K}$ est un corps quelconque alors toute extension séparable finie est simple.

Démonstration. Nous ne donnons la preuve que dans le cas où $\mathbb{K}$ est fini. Dans ce cas nous savons par la proposition 20.33 que le groupe $\mathbb{K}^*$ est cyclique. Si de plus $\mathbb{L}$ est une extension finie alors $\mathbb{L}$ est fini en tant qu'ensemble. Par conséquent $\mathbb{L}^*$ est un groupe cyclique. Si α 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. □

Proposition 20.52.
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 $p^n - 1$.

Lemme 20.53.
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 :

$$(\alpha + \beta)^p = \sum_k \binom{k}{p} \alpha^k \beta^{p-k} \tag{20.192}$$

où les coefficients binomiaux sont dans $\mathbb{F}_p$ et donc nuls pour les k différents de p et de 0. □

Cette proposition est encore vraie avec $\alpha, \beta \in \mathbb{F}_{p^n}$ et $(\alpha + \beta)^{p^n}$.

Lemme 20.54.
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 $\{\alpha^i$ tel que $i = 0, \dots, k-1\}$.

20. Définition 6.90.

Nous serions donc intéressé 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.

Théorème 20.55 (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 une polynôme irréductible $P \in \mathbb{F}_p[X]$ de degré n tel que

$$\begin{aligned} \phi\colon \mathbb{F}_p[X]/(P) &\to \mathbb{K} \\ \bar{X} &\mapsto \alpha \end{aligned} \tag{20.193}$$

soit un isomorphisme de corps.

Soit α et P choisis pour avoir les propriétés citées plus haut. Alors nous avons les propriétés suivantes.

(1) P est primitif[21]*.*

(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 20.33. Si α un générateur de $\mathbb{K}^*$ alors

$$\mathbb{K} = \mathbb{F}_p[\alpha]. \tag{20.194}$$

Soit ℓ le plus grand entier tel que l'ensemble

$$\{1, \alpha, \cdots, \alpha^{\ell-1}\} \subset \mathbb{K} \tag{20.195}$$

soit libre. Pour rappel $\mathbb{K}$ est une espace vectoriel sur $\mathbb{F}_p$. Il existe des $a_i \in \mathbb{F}_p$ tels que

$$\alpha^\ell + a_{\ell-1}\alpha^{\ell-1} + \cdots + a_0 = 0. \tag{20.196}$$

De façon équivalente, il existe un polynôme unitaire $P \in \mathbb{F}_p[X]$ de degré ℓ tel que $P(\alpha) = 0$. Étant donné que α est générateur de $\mathbb{K}$,

$$\mathbb{K} = \text{Span}\{1, \alpha, \dots, \alpha^{\ell-1}\} \tag{20.197}$$

parce que $\mathbb{K}$ est généré par les puissances de α alors que les puissances de α 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 ℓ ; par conséquent

$$\text{Card}(\mathbb{K}) = p^n = q \tag{20.198}$$

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

$$\{\alpha^{\ell-1}, \dots, 1\} \tag{20.199}$$

est libre.

Montrons que l'application

$$\begin{aligned} \phi\colon \mathbb{F}_p[X]/(P) &\to \mathbb{K} \\ \bar{X} &\mapsto \alpha \end{aligned} \tag{20.200}$$

21. Définition 3.168.

est un isomorphisme. Pour l'injectivité, deux éléments $Q_1, Q_2 \in \mathbb{F}_p[X]/(P)$ s'écrivent

$$Q_1 = \sum_{k=0}^{n-1} a_k \bar{X}^k \tag{20.201a}$$

$$Q_2 = \sum_{k=0}^{n-1} b_k \bar{X}^k. \tag{20.201b}$$

Dans ce cas si $\phi(Q_1) = \phi(Q_2)$ alors

$$\phi(Q_1) = \sum_{k=0}^{n-1} a_k \alpha^k = \phi(Q_2) = \sum_{k=0}^{n-1} b_k \alpha^k. \tag{20.202}$$

Mais l'ensemble $\{1, \alpha, \ldots, \alpha^{n-1}\}$ étant libre sur $\mathbb{F}_p$, cela implique $a_k = b_k$. La surjectivité de ϕ provient du fait que α 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 $\alpha \mapsto \bar{X}$ soit un isomorphisme entre $\mathbb{K}$ et $\mathbb{F}_p[X]/(P)$.

Le polynôme P est primitif parce que α est d'ordre p^n dans $\mathbb{K}$ alors que $\bar{X} \mapsto \alpha$ est un isomorphisme. Par conséquent $\bar{X}$ est d'ordre p^n dans $\mathbb{F}_p[X]/P$.

Nous commençons par prouver que l'ensemble

$$\{\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}}\} \tag{20.203}$$

est l'ensemble des racines distinctes de P. Pour cela nous posons

$$P(X) = \sum_{k=0}^{n} a_k X^k \tag{20.204}$$

avec $a_k \in \mathbb{F}_p$. D'abord α est une racine de P. En effet

$$P(\bar{X}) = \sum_k a_k \bar{X}^k = 0 \tag{20.205}$$

parce que cette somme est calculée dans $\mathbb{F}_p[X]/(P)$. En appliquant l'isomorphisme ϕ à l'égalité (20.205) nous trouvons

$$0 = \phi\big(P(\bar{X})\big) = \sum_k a_k \phi(\bar{X}^k) = \sum_k a_k \alpha^k. \tag{20.206}$$

Donc α 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 (20.203). Étant donné que pour tout x dans $\mathbb{F}_p$ nous avons $x^p = x$, nous avons aussi

$$P(X^p) = \sum_k a_k (X^p)^k = \sum_k a_k^p (X^p)^k = \sum_k (a_k X^k)^p \tag{20.207}$$

alors que nous savons que $x \mapsto x^p$ est un automorphisme de $\mathbb{F}_p$ par la proposition 3.59. Par conséquent

$$P(X^p) = \sum_k (a_k X^k)^p = \left(\sum_k a_k X^k \right)^p = P(X)^p. \tag{20.208}$$

Nous avons montré que si β est une racine de P, alors β^p est également une racine de P. Nous savons déjà que α est une racine de P, et que α est également générateur de $\mathbb{K}$, c'est-à-dire que α est d'ordre $q - 1$. Les puissances

$$\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}} \tag{20.209}$$

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

$$\{\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}}\} \tag{20.210}$$

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 β une racine de P ; il s'écrit $\beta = \alpha^k$ pour un certain k. Étant donné que $\alpha^{q-1} = e = \alpha^{p^n-1}$,

$$\begin{aligned}\beta^q &= (\alpha^{p^n})^k && (20.211a)\\ &= \left(\alpha^{p^n-1}\alpha\right)^k && (20.211b)\\ &= \left(\alpha^{q-1}\alpha\right)^k && (20.211c)\\ &= \alpha^k && (20.211d)\\ &= \beta. && (20.211e)\end{aligned}$$

Cela signifie que $\beta^q = \beta$ et donc que β est racine de $X^q - X$. □

Corollaire 20.56.
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 (20.193). Soit 1_q la classe du polynôme 1 modulo P, nous considérons le morphisme

$$\begin{aligned}\mu\colon \mathbb{Z} &\to \mathbb{F}_q\\ n &\mapsto n1_q.\end{aligned} \tag{20.212}$$

Le noyau de cette application est $\ker\mu = \mathbb{Z}_p$ parce que $p1_q = 0$, les coefficients étant à comprendre dans $\mathbb{F}_p$. □

Définition 20.57.
Soit 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 α parcourent tout le groupe multiplicatif $(\mathbb{F}_p[X]/P)^*$.*

Lemme 20.58.
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 α^p est également une racine parce que si $P = \sum_k a_k X^k$,

$$P(\alpha^p) = \sum_k (a_k\alpha^k)^p = \Big(\sum_k a_k\alpha^k\Big)^p = 0 \tag{20.213}$$

où nous avons utilisé le fait que $a_k^p = a_k$ étant donné que $a_k \in \mathbb{F}_p$. Par hypothèse α est une racine primitive ; cela implique que les éléments $\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^n-1}$ 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 α est une puissance de β. Étant donné que $(\mathbb{F}_p[X]/P)^*$ est un groupe à $p^n - 1$ éléments, le corollaire 2.32 indique que $\alpha^{p^n} = \alpha$. En particulier avec $r = p^{n-k}$ nous avons

$$\beta^r = \alpha^{rp^k} = \alpha^{p^n} = \alpha. \tag{20.214}$$

Par suite toutes les puissances de α sont des puissances de β, ce qui implique que β est générateur du groupe cyclique $(\mathbb{F}_p[X]/P)^*$. □

Lemme 20.59.
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 20.60.
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.

Proposition 20.61.
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 20.55. L'élément a est en particulier une racine de $X^q - X$. Par ailleurs P divise $X^q - X$ par le lemme 20.59.

Nous avons aussi

$$X^q - X = \prod_{b\in\mathbb{L}}(X - b) \tag{20.215}$$

par la proposition 20.35. 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 avons que Q divise P, mais P étant irréductible et unitaire nous avons immédiatement $P = Q$. En particulier nous avons

$$\mathbb{F}_p[X]/P \simeq \mathbb{F}_p[X]/Q \simeq \mathbb{K}. \tag{20.216}$$

Nous montrons maintenant que $\mathbb{F}_p[X]/Q \simeq \mathbb{L}$ par l'application

$$\begin{aligned}\phi\colon \mathbb{F}_p[X]/Q &\to \mathbb{L}\\ \bar{X} &\mapsto b\end{aligned} \tag{20.217}$$

qui se prolonge en $R(\bar{X}) \mapsto R(b)$ 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. □

20.5.7 Construction de $\mathbb{F}_q$

Soit p un nombre premier et $n \in \mathbb{N}$. Nous souhaitons construire un corps à $q = p^n$ éléments. Nous savons déjà que ce corps est unique (théorème 20.32) et que nous le notons $\mathbb{F}_q$. Le théorème 20.55 nous incite à l'écrire sous la forme

$$\mathbb{F}_q = \mathbb{F}_p[X]/(P) \tag{20.218}$$

pour un certain polynôme irréductible $P \in \mathbb{F}_p[X]$.

20.5.7.1 La version du faignant

Nous pouvons construire le corps à 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.

Proposition 20.62.
Soit P une 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) $\alpha = \bar{X}$ est une racine de P dans $\mathbb{K}$.

(3) $\mathbb{K} = \mathbb{F}_p[\alpha]$.

Démonstration. (1) En vertu du corollaire 6.37, $\mathbb{K}$ 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 $P(\bar{X}) = 0$ 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 $\bar{X}$.

□

20.5.7.2 La version plus élaborée

Construire $\mathbb{F}_q$ comme quotient de $\mathbb{F}_p[X]$ par un polynôme irréductible quelconque ne donne pas d'informations sur les générateurs de $\mathbb{F}_q^*$, et en particulier il n'est pas toujours vrai que $\bar{X}$ est générateur.

Exemple 20.63
Construisons $\mathbb{F}_4$. Le polynôme X^2+X+1 est irréductible dans $\mathbb{F}_2$ parce qu'il n'a pas de racines (c'est vite vu : dans $\mathbb{F}_2$ il n'y a que deux candidats). Donc $\mathbb{F}_4 = \mathbb{F}_2[X]/(X^2+X+1)$. △

Remarque 20.64.
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.

Exemple 20.65
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 a avoir des racines génératrices sont donc au nombre de 3 :

$$P_1 = X^4 + X + 1 \tag{20.219a}$$
$$P_2 = X^4 + X^3 + 1 \tag{20.219b}$$
$$P_3 = X^4 + X^3 + X^2 + X + 1. \tag{20.219c}$$

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

$$X \tag{20.220a}$$
$$X^2 \tag{20.220b}$$
$$X^3 \tag{20.220c}$$
$$X^4 = X^3 + X^2 + X + 1 \tag{20.220d}$$
$$1. \tag{20.220e}$$

La classe de X dans $\mathbb{F}_2[X]/P_3$ n'est donc pas génératrice du groupe $(\mathbb{F}_2[X]/P_3)^*$.

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

$$X \tag{20.221a}$$
$$X^2 \tag{20.221b}$$
$$X^3 \tag{20.221c}$$
$$X+1 \tag{20.221d}$$
$$X^2+X \tag{20.221e}$$
$$X^3+X^2 \tag{20.221f}$$
$$X+1+X^3 \tag{20.221g}$$
$$X^2+1 \tag{20.221h}$$
$$X^3+X \tag{20.221i}$$
$$X+1+X^2 \tag{20.221j}$$
$$X^2+X+X^3 \tag{20.221k}$$
$$X^3+X^2+X+1 \tag{20.221l}$$
$$1+X^2+X^2 \tag{20.221m}$$
$$1+X^3 \tag{20.221n}$$
$$1 \tag{20.221o}$$

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. △

Proposition 20.66 ([1])**.**
Soit P un polynôme irréductible unitaire primitif dans $\mathbb{F}_p[X]$. Nous considérons $\mathbb{K}=\mathbb{F}_p[X]/P$ et $\alpha=\bar{X}\in\mathbb{K}$. Alors

(1) Les racines de P sont $\{\alpha,\alpha^p,\ldots,\alpha^{p^{n-1}}\}$ et $\alpha^q=\alpha$.

(2) P est le polynôme minimal de α.

(3) P est scindé dans $\mathbb{K}$.

(4) P divise X^q-X dans $\mathbb{K}$.

(5) La famille $\{1,\alpha,\alpha^2,\ldots,\alpha^{n-1}\}$ est une base de $\mathbb{K}$ en tant qu'espace vectoriel sur $\mathbb{F}_p$.

(6) En tant qu'ensemble,

$$\mathbb{F}_q=\{0,\alpha,\alpha^2,\alpha^3,\ldots,\alpha^{q-1}\}, \tag{20.222}$$

et les α^k sont distincts pour $k=1,\ldots,q-1$.

Démonstration. La plupart des assertions sont des corollaires ou des paraphrases de résultats contenus dans les propositions précédentes.

(1) L'assertion à propos des racines de P est contenue dans le lemme 20.58. D'autre part le groupe $(\mathbb{F}_p[X]/P)^*$ est cyclique d'ordre $q-1$. Par conséquent le corollaire 2.32 indique que $\alpha^{q-1}=1$ et donc $\alpha^q=\alpha$.

(2) Soit $\tilde{P}$ un polynôme annulateur de α. Nous voyons que si β est racine de $\tilde{P}$ alors β^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 20.35 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 $\{1,\alpha,\alpha^2,\ldots,\alpha^{n-1}\}$ serait un polynôme annulateur de degré $n-1$ de α. 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 $k, l \leqslant q$ alors nous avons $\alpha^r = 1$ avec $r = l - k < q$, ce qui contredirait la primitivité de P. Les éléments $0, \alpha, \ldots, \alpha^{q-1}$ étant distincts et au nombre de q, ils forment tout l'ensemble $\mathbb{F}_q$.

□

20.5.8 Exemple : étude de $\mathbb{F}_{16}$

Dans cette sous section nous voulons construire $\mathbb{F}_{16}$. Nous considérons donc $p = 2$ et $n = 4$. Des polynôme irréductibles de degré 4 dans $\mathbb{F}_2[X]$ ne sont pas très difficiles à trouver. Par exemple $X^4 + X^3 + X^2 + X + 1$, plus généralement un polynôme contenant un nombre impair de termes non nuls dont le terme indépendant.

Les polynômes primitifs par contre doivent être trouvés parmi les diviseurs irréductibles de $X^{15} - 1$. Montrons que

$$P = X^4 + X^3 + 1 \tag{20.223}$$

est primitif. Nous posons $\omega = \bar{X} \in \mathbb{F}_2[X]/P$. L'ordre de ω dans le groupe $(\mathbb{F}_2[X]/P)^*$ doit être un diviseur de 15 et donc 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 ω n'est pas 5 non plus :

$$\omega^5 = \omega^4\omega = (\omega^3 + 1)\omega = \omega^4 + \omega = \omega^3 + \omega + 1 \neq 1. \tag{20.224}$$

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 ω^8. En effet si β est une racine de P, alors β^2 est une racine en vertu de

$$P(\beta^2) = (\beta^2)^4 + (\beta^2)^3 + 1 = (\beta^4)^2 + (\beta^3)^2 + 1^2 = (\beta^4 + \beta^3 + 1)^2 = 0. \tag{20.225}$$

Ici nous avons implicitement utilisé le lemme 20.53. D'autre part P ne peut pas avoir plus de 4 racines.

Proposition 20.67.
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 ω 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

$$f_1 = \omega \tag{20.226a}$$
$$f_2 = \omega^2 \tag{20.226b}$$
$$f_3 = \omega^3 + 1 \tag{20.226c}$$
$$f_4 = \omega^3 + \omega^2 + \omega. \tag{20.226d}$$

Ensuite nous montrons que les vecteurs e_i peuvent être construits comme combinaisons linéaires des vecteurs f_j :

$$f_1 + f_2 + f_3 + f_4 = e_0 \tag{20.227a}$$
$$f_1 = e_1 \tag{20.227b}$$
$$f_2 = e_2 \tag{20.227c}$$
$$f_1 + f_2 + f_4 = e_3. \tag{20.227d}$$

Les quatre vecteurs f_j forment donc bien un base parce qu'ils sont générateurs d'un espace de dimension 4. □

Exemple 20.68

(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_γ est soit un élement de B_γ 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 ω ; nous posons $a = \omega^l$. Nous cherchons x sous la forme $x = \omega^k$. L'équation à résoudre pour k est

$$\omega^{5k} = \omega^l \tag{20.228}$$

où l est donné. Cette équation revient à

$$5k = l \mod 15. \tag{20.229}$$

Si l n'est pas un multiple de 5, alors il n'y a pas de solutions. Il n'y a des solutions uniquement pour $l = 0, 5, 10$ et elles sont :

$$k = \begin{cases} 3, 6, 9, 12 & \text{si } l{=}0 \\ 1 & \text{si } l{=}5 \\ 2 & \text{si } l{=}10 \end{cases} \tag{20.230}$$

(2) Nous cherchons γ sous la forme $\gamma = \omega^k$. Parmi les nombreuses contraintes liées à l'énoncé nous devons avoir

$$\gamma^5 = 1, \gamma, \gamma^2, \gamma^4, \gamma^8. \tag{20.231}$$

Les possibilités $\gamma^5 = \gamma, \gamma^2, \gamma^4, \gamma^5$ ne sont pas bonnes parce qu'elles impliqueraient que B_γ n'est pas une base. Reste à explorer $\gamma^5 = 1$.
Étant donné le premier point nous restons avec les possibilités

$$\gamma = 1, \omega^3, \omega^6, \omega^9, \omega^{12}. \tag{20.232}$$

Évidemment $\gamma = 1$ ne produit pas une base. Avec $\gamma = \omega^3$ nous trouvons

$$B_\gamma = \{\omega^3, \omega^6, \omega^{12}, \omega^{24}\} = \{\omega^3, \omega^6, \omega^{12}, \omega^9\} \tag{20.233}$$

où nous avons utilisé le fait que $\omega^k = \omega^{k \mod 15}$. En utilisant le fait que $\omega^4 = \omega^3 + 1$ nous trouvons

$$\omega^5 = \omega^3 + \omega + 1 \tag{20.234a}$$
$$\omega^6 = \omega^3 + \omega^2 + \omega + 1 \tag{20.234b}$$
$$\omega^9 = \omega^2 + 1 \tag{20.234c}$$
$$\omega^{12} = \omega + 1. \tag{20.234d}$$

L'ensemble B_γ est alors formé des éléments

$$f_1 = \omega^3 \tag{20.235a}$$
$$f_2 = \omega^3 + \omega^2 + \omega + 1 \tag{20.235b}$$
$$f_3 = \omega + 1 \tag{20.235c}$$
$$f_4 = \omega^2 + 1. \tag{20.235d}$$

Il est assez simple de vérifier que cela est une base en remarquant que $f_1 + f_2 + f_2 + f_4 = 1$. Les possibilités $\gamma = \omega^6, \omega^9, \omega^{12}$ produisent les mêmes ensemble B_γ.

△

20.5.9 Polynômes irréductibles sur $\mathbb{F}_q$

Définition 20.69.
*La **fonction de Möbius** est la fonction $\mu\colon \mathbb{N}^* \to \{-1,0,1\}$ définie par*

$$\mu(n) = \begin{cases} 0 & \text{si } n \text{ est divisible par un carré différent de } 1,\\ 1 & \text{si } n \text{ est le produit d'un nombre pair de nombres premiers distincts,}\\ -1 & \text{si } n \text{ est le produit d'un nombre impair de nombres premiers distincts,}\end{cases} \tag{20.236}$$

Proposition 20.70 ([343]).
Si m et n sont strictement positifs et premiers entre eux, alors

$$\mu(mn) = \mu(m)\mu(n). \tag{20.237}$$

De plus nous avons

$$\sum_{d\,|\,n} \mu(d) = \begin{cases} 1 & \text{si } n = 1\\ 0 & \text{si } n > 1\end{cases} \tag{20.238}$$

Proposition 20.71 (Formule d'inversion de Möbius[343]).
Soient $f, g\colon \mathbb{N} \to \mathbb{C}$ telles que pour tout $n \geqslant 1$,

$$g(n) = \sum_{d\,|\,n} f(d). \tag{20.239}$$

Alors

$$f(n) = \sum_{d\,|\,n} \mu\left(\frac{n}{d}\right) g(d) \tag{20.240}$$

où μ est la fonction de Möbius pour tout $n \geqslant 1$.

Lemme 20.72 ([344]).
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 $a, b \in \mathbb{K} \subset \mathbb{L}$ tels que [22] $aP + bQ = 1$. Cette dernière égalité est encore valable dans $\mathbb{L}$ et donc rend impossible l'existence d'une racine commune. □

Proposition 20.73 ([344, 43]).
Soit p un nombre premier, $n \geqslant 1$ et $r \in \mathbb{N}^*$. Nous notons $q = p^r$, $A(n,q)$, l'ensemble des polynômes unitaires irréductibles de degré n sur $\mathbb{F}_q$. Nous notons aussi $I(n,q) = \operatorname{Card}\big(A(n,q)\big)$. Alors :

(1) Le polynôme $X^{q^n} - X$ se décompose en irréductibles de la façon suivante :

$$X^{q^n} - X = \prod_{d\,|\,n} \prod_{p \in A(d,q)} P. \tag{20.241}$$

(2) Le nombre d'irréductibles est donné par

$$I(n,q) = \frac{1}{n} \sum_{d\,|\,n} \mu\left(\frac{n}{d}\right) q^d \tag{20.242}$$

où μ est la fonction de Möbius (définition 20.69).

(3) Nous avons l'équivalence de suite

$$I(n,q) \sim_{n\to\infty} \frac{q^n}{n}. \tag{20.243}$$

22. Théorème de Bézout, 6.40.

Démonstration. (1) Soit un diviseur d de n et $P \in A(d,q)$. Montrons que P divise $X^{q^n}-X$. Nous considérons la 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 $\deg(P)$ a coefficients dans $\mathbb{F}_q$. Ce corps possède donc q^d éléments et est isomorphe à $\mathbb{F}_{q^d}$ par la proposition 20.61. 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 20.35.

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 α est aussi une racine de $X^{q^n}-X$. En effet en utilisant le fait que $\alpha^{q^d}=\alpha$, nous avons

$$\alpha^{q^n} = \alpha^{q^{kd}} = \alpha^{q^d q^{(k-1)d}} = \left(\alpha^{q^d}\right)^{q^{(k-1)d}} = \alpha^{q^{(k-1)d}}, \tag{20.244}$$

donc par récurrence, on a encore $\alpha^{q^n}=\alpha$, et α est racine de $X^{q^n}-X$. Vu que P est irréductible, le lemme 20.72 nous indique que P divise $X^{q^n}-X$. Nous en déduisons 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 $X^{q^n}-X$:

$$\prod_{d\,|\,n}\prod_{P\in A(d,q)} P \mid X^{q^n}-X. \tag{20.245}$$

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 $d \geqslant 1$. Nous posons encore $\mathbb{K} = \mathbb{F}_q[X]/(P)$ et nous utilisons la propriété de multiplication sur les degrés (proposition 6.57) :

$$[\mathbb{F}_{q^n} : \mathbb{K}][\mathbb{K} : \mathbb{F}_q] = [\mathbb{F}_{q^n} : \mathbb{F}_q] = n, \tag{20.246}$$

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 20.35), dans sa décomposition en irréductibles sur $\mathbb{F}_q$, il n'a pas de facteurs carrés ; 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 $\prod_{d\,|\,n}\prod_{P\in A(d,q)} P$ et donc $X^{q^n}-X$ divise ce gros produit :

$$X^{q^n}-X \mid \prod_{d\,|\,n}\prod_{P\in A(d,q)} P. \tag{20.247}$$

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 :

$$q^n = \sum_{d\,|\,n} d\,\mathrm{Card}\left(A(d,q)\right) = \sum_{d\,|\,n} dI(d,q). \tag{20.248}$$

Nous pouvons utiliser la formule d'inversion de Möbius (proposition 20.71) pour les fonctions $g(n)=q^n$ et $f(n)=dI(n,q)$. Nous écrivons alors

$$f(n) = \sum_{d\,|\,n} \mu\left(\frac{n}{d}\right) q^d, \tag{20.249}$$

ou encore

$$I(n,q) = \frac{1}{n}\sum_{d\,|\,n} \mu\left(\frac{n}{d}\right) q^d, \tag{20.250}$$

ce qu'il fallait.

(3) Nous posons

$$r_n = \sum_{\substack{d|n \\ d<n}} \mu\left(\frac{n}{d}\right) q^d, \tag{20.251}$$

mais sachant que les diviseurs de n, outre n lui-même, sont tous plus petit ou égal à $n/2$ et qu'en valeur absolue, la fonction de Möbius est toujours plus petite ou égale à [23] 1,

$$|r_n| \leqslant \sum_{d=1}^{\lfloor n/2 \rfloor} q^d = \frac{q - q^{\lfloor n/2 \rfloor}}{1-q} = q\frac{q^{\lfloor n/2 \rfloor} - 1}{q-1} \leqslant \frac{q^{\lfloor n/2 \rfloor + 1}}{q-1}. \tag{20.252}$$

D'autre part en reprenant la formule déjà prouvée,

$$I(n,q) = \frac{1}{n}\sum_{d|n} \mu\left(\frac{n}{d}\right) q^d = \frac{1}{n}\left(r_n + \mu\left(\frac{n}{n}\right) q^n\right) = \frac{r_n + q^n}{n}. \tag{20.253}$$

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 $n \to \infty$:

$$\frac{q^n + r_n}{n} \sim_{n\to\infty} \frac{q^n}{n}. \tag{20.254}$$

□

20.5.10 Matrices

Proposition 20.74.
Nous avons

$$|\mathrm{GL}(n, \mathbb{F}_p)| = (p^n - 1)(p^n - p)\dots(p^n - p^{n-1}). \tag{20.255}$$

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. □

Lemme 20.75.
Soit $\mathbb{K}$ un corps fini autre que $\mathbb{F}_2$ [24], soit un groupe abélien M et un morphisme $\varphi\colon \mathrm{GL}(n, \mathbb{K}) \to M$. Alors il existe un unique morphisme $\delta\colon \mathbb{K}^* \to M$ 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 $D\big(\mathrm{GL}(n, \mathbb{K})\big)$ 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

$$\tilde{\varphi}\colon \frac{\mathrm{GL}(n, \mathbb{K})}{\mathrm{SL}(n, \mathbb{K})} \to M \tag{20.256}$$

vérifiant $\varphi = \tilde{\varphi} \circ \pi$ où π est la projection.

Nous pouvons faire la même chose avec l'application

$$\det\colon \mathrm{GL}(n, \mathbb{K}) \to \mathbb{K}^* \tag{20.257}$$

qui est un morphisme de groupes dont le noyau est $\mathrm{SL}(n, \mathbb{K})$. Cela nous donne une application

$$\tilde{\det}\colon \frac{\mathrm{GL}(n, \mathbb{K})}{\mathrm{SL}(n, \mathbb{K})} \to \mathbb{K}^* \tag{20.258}$$

23. Dans [43], ma dernière inégalité arrive comme une égalité.
24. Je ne comprends pas très bien à quel moment joue cette hypothèse.

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

$$\varphi = \tilde{\varphi} \circ \tilde{\det}^{-1} \circ \tilde{\det} \circ \pi. \tag{20.259}$$

État 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 $\delta' \colon \mathbb{K}^* \to M$ 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 $\mathbb{K}^*$, nous avons $\delta' = \delta$. □

Théorème 20.76.
Soit $p \geqslant 3$ un nombre premier et V, un $\mathbb{F}_p$-espace vectoriel de dimension finie n. Pour tout $u \in \mathrm{GL}(V)$ nous avons

$$\epsilon(u) = \left(\frac{\det(u)}{p} \right). \tag{20.260}$$

Ici ϵ est la signature de u vue comme une permutation des éléments de $\mathbb{F}_p$.

Démonstration. Commençons par prouver que

$$\epsilon \colon \mathrm{GL}(V) \to \{-1, 1\}. \tag{20.261}$$

est un morphisme. Si nous notons $\bar{u} \in S(V)$ l'élément du groupe symétrique correspondant à la matrice $u \in \mathrm{GL}(V)$, alors nous avons $\overline{uv} = \bar{u} \circ \bar{v}$, et la signature étant un homomorphisme (proposition 2.73),

$$\epsilon(uv) = \epsilon(\bar{u} \circ \bar{v}) = \epsilon(\bar{u})\epsilon(\bar{v}). \tag{20.262}$$

Par ailleurs $\{-1, 1\}$ est abélien, donc le lemme 20.75 s'applique et nous pouvons considérer un morphisme $\delta \colon \mathbb{F}_p^* \to \{-1, 1\}$ tel que $\epsilon = \delta \circ \det$.

Nous allons utiliser le lemme 20.43 pour montrer que δ est le symbole de Legendre. Pour cela il nous faudrait trouver un $x \in \mathbb{F}_p^*$ 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

$$\delta(x) = (\delta \circ \det)(u) = \epsilon(u), \tag{20.263}$$

et notre problème revient à trouver une matrice $u \in \mathrm{GL}(V)$ dont la permutation associée soit de signature -1.

Soit $n = \dim V$; en conséquence de la proposition 20.62(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 V. Étant donné que $\mathbb{F}_q$ est un corps fini, nous savons que $\mathbb{F}_q^*$ est un groupe cyclique à $q-1$ éléments. Soit y, un générateur de $\mathbb{F}_q^*$ et l'application

$$\begin{aligned} \beta \colon \mathbb{F}_q &\to \mathbb{F}_q \\ x &\mapsto yx. \end{aligned} \tag{20.264}$$

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 β fixe zéro et à part zéro, agit comme le cycle

$$(1, y, y^2, \dots, y^{q-2}). \tag{20.265}$$

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 $p \geqslant 3$) et par conséquent, l'application β est de signature -1. □

20.6 Constructions à la règle et au compas

Définition 20.77 ([345]).
Soit E une partie de $\mathbb{R}^2$. Un point de $\mathbb{R}^2$ est constructible en une étape à partir de E s'il est une 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 : $C_{n+1}(E) = C_1\big(C_n(E)\big)$. Enfin un point de $\mathbb{R}^2$ est **constructible** à partir de E s'il appartient à*

$$C(E) = \bigcup_{n=1}^{\infty} C_n(E). \tag{20.266}$$

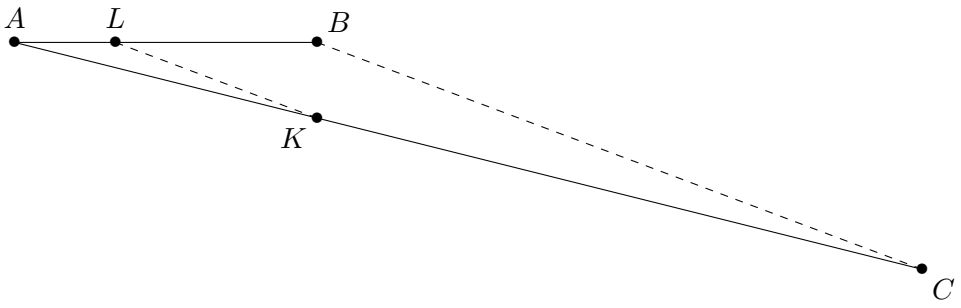
*Un réel est **constructible** s'il est l'abscisse d'un point constructible.*

Pour toute la suite nous allons considérer les points et réels constructibles à partir de l'ensemble $E = \{(0,0),(0,1)\}$.

20.6.1 Quelques constructions

Proposition 20.78 ([346]).
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 $n \in \mathbb{N}^*$.

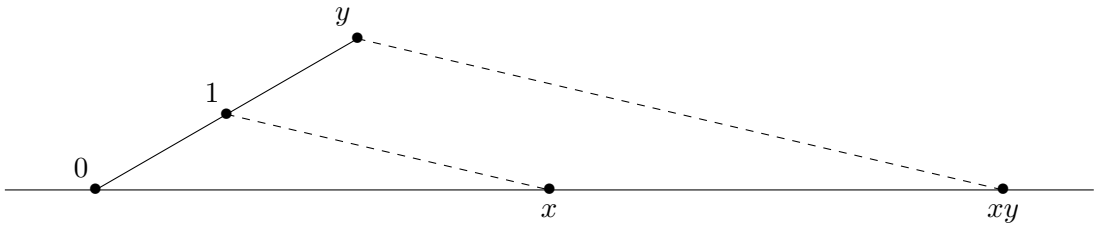


La méthode pour construire le nombre $1/n$ est la suivante. Soit $[AB]$ un segment de longueur 1 (par exemple $A = (0,0)$ et $B = (1,0)$) et un point C tel que $[AC]$ ait une longueur n. Nous plaçons sur $[AC]$ le point K situé à une distance 1 de A en pointant le compas en A et en traçant le cercle de rayon $[AB]$.

La droite passant par K et parallèle à (BC) coupe $[AB]$ en un point L. Maintenant $[AL]$ a une longueur $1/n$ par le théorème de Thalès.

□

Exemple 20.79(Multiplication à la règle et au compas[345])
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. △

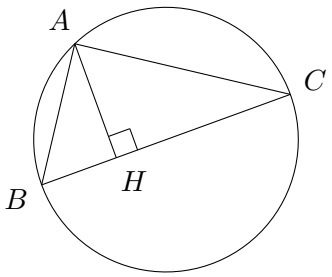
Exemple 20.80(Racine carré à la règle et au compas[345])

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, 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 :

$$\frac{AB}{BH}=\frac{BC}{AB}, \tag{20.267}$$

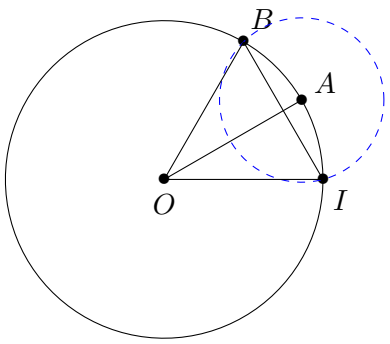
ce qui donne $AB^2 = BC \times BH = x$ (BC et BH valent respectivement 1 et x ou le contraire).



△

Exemple 20.81(Duplication d'un angle)
Si un angle α est constructible, nous allons construire les angles 2α, 3α, 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 3α.

△

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.

Exemple 20.82(Addition d'angles)
Quitte à soustraire ou additionner un certain nombre de fois 90°, nous supposons que les deux angles donnés sont entre 0° et 90°.

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 tout deux égaux à α.

Les angles $\widehat{K_1OA}$ et $\widehat{K_2OA}$ sont égaux à $\beta - \alpha$ et $\beta + \alpha$. △

20.6.2 Nombres constructibles

Théorème 20.83 (Wantzel[347]).
Le réel a est constructible si et seulement s'il existe une suite finie de corps $\mathbb{L}_i$ tels que

(1) $\mathbb{L}_0 = \mathbb{Q}$,

(2) $\mathbb{L}_{i+1}$ *est un extension quadratique*[25] *de* $\mathbb{L}_i$

(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.

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

$$\begin{cases} ax + by + c = 0 & \text{(20.268a)} \\ a'x + b'y + c' = 0, & \text{(20.268b)} \end{cases}$$

dont les solutions sont encore dans $\mathbb{K}$.

Intersection droite-cerle L'équation d'un cercle est de la forme

$$(x - u)^2 + (y - v)^2 = r^2 \tag{20.269}$$

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

$$x^2 + y^2 + ux + vy + t = 0 \tag{20.270}$$

avec $u, v, t \in \mathbb{K}$. Il s'agit de voir où sont les solutions du système

$$\begin{cases} ax + by + c = 0 & \text{(20.271a)} \\ x^2 + y^2 + ux + vy + t = 0. & \text{(20.271b)} \end{cases}$$

Si $a \neq 0$ alors nous pouvons faire la substitution $x = -(c + by)/a$ et obtenir l'équation suivante pour y :

$$\left(\frac{-c - by}{a}\right)^2 + y^2 + u\left(\frac{-c - by}{a}\right) + vy + t = 0. \tag{20.272}$$

Cela est de la forme $P(y) = 0$ où P est un polynôme du second degré à coefficients dans $\mathbb{K}$.

25. C'est-à-dire une extension finie de degré 2.

Si $a = 0$ nous substituons y au lieu de x et le résultat est le même.

C'est le moment de relire la proposition 6.85 qui nous assure que si le réel α 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 α est de degré 2 :

$$\big[\mathbb{K}[\alpha] : \mathbb{K}\big] = 2. \tag{20.273}$$

De plus $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$.

Intersection cercle-cercle Le système

$$\begin{cases} x^2 + y^2 + ux + vy + y = 0 & \text{(20.274a)} \\ x^2 + y^2 + ax + by + c = 0 & \text{(20.274b)} \end{cases}$$

est équivalent au système (en substituant la seconde équation par la différence entre les deux)

$$\begin{cases} x^2 + y^2 + ux + vy + y = 0 & \text{(20.275a)} \\ (u-a)x + (v-b)y + t - c = 0 & \text{(20.275b)} \end{cases}$$

qui est à nouveau une intersection entre un cercle et une droite.

Passons à la conclusion. Si α 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 pour chaque réel constructible en $m-1$ étapes possède sa pile d'extension quadratiques en partant de $\mathbb{Q}$. Le nombre α vérifie donc

$$a\alpha^2 + b\alpha + c = 0 \tag{20.276}$$

avec a, b, c trois réels constructibles en $m-1$ étapes. L'hypothèse de récurrence donne donc des piles d'extensions

$$\begin{array}{ccc} \mathbb{L}_0 = \mathbb{Q} & \mathbb{L}'_0 = \mathbb{Q} & \mathbb{L}''_0 = \mathbb{Q} \\ L_{i+1} = \mathbb{L}_i(a_i) & \mathbb{L}'_{i+1} = \mathbb{L}'_i(b_i) & \mathbb{L}''_{i+1} = \mathbb{L}''_i(c_i) \\ a \in \mathbb{L}_n & b \in \mathbb{L}'_{n'} & c \in \mathbb{L}''_{n''}. \end{array} \tag{20.277}$$

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

$$\begin{aligned} \mathbb{M}_0 &= \mathbb{Q} \\ \mathbb{M}_{i+1} &= \mathbb{M}_i(\alpha_i) \\ a, b, c &\in \mathbb{M}_n. \end{aligned} \tag{20.278}$$

Ici le n n'est pas spécialement le même que celui plus haut. Maintenant, α est dans un 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 α est constructible. Nous y allons par récurrence : si $n = 0$ alors α est rationnel et il est constructible par la proposition 20.78.

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. Vu que α est solution d'une équation de degré 2 à coefficients dans $\mathbb{L}_i$, nous avons

$$a\alpha^2 + b\alpha + c = 0 \tag{20.279}$$

pour certains $a, b, c \in \mathbb{L}_i$ et donc

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{20.280}$$

Les exemples 20.79 et 20.80 montrent que les produits et les racines carrées de nombres constructibles sont constructibles. Donc α est constructible.

□

20.6.3 Polygones constructibles

Définition 20.84.
*Un angle α est **constructible** si le nombre $\cos(\alpha)$ est constructible.*

La raison est que si qu'il suffit de prendre la perpendiculaire à l'axe horizontal

Lemme 20.85 ([43]).
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. **Sens direct** Il suffit de pouvoir multiplier un angle par un entier, ce qui est fait dans l'exemple 20.81.

Sens réciproque Le théorème de Bézout 3.13 nous donne $a, b \in \mathbb{Z}$ tels que $an + bm = 1$. Cela donne immédiatement

$$\frac{1}{mn} = \frac{a}{m} + \frac{b}{n}, \tag{20.281}$$

et donc

$$\frac{2\pi}{mn} = a\frac{2\pi}{m} + b\frac{2\pi}{n}, \tag{20.282}$$

ce qui fait que l'angle $2\pi/mn$ est une combinaison entière d'angles constructibles. Il est donc constructible par l'exemple 20.82.

□

Lemme 20.86.
Soit un entier $n \geqslant 3$ ayant

$$n = \prod_{i=1}^{k} p_i^{\alpha_i} \tag{20.283}$$

comme décomposition en facteurs premiers. Les 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

$$\frac{2\pi}{\prod_{i=1}^{k} p_i^{\alpha_i}} \tag{20.284}$$

est constructible. Par récurrence sur le lemme 20.85, cet angle est constructible si et seulement si les angles $\frac{2\pi}{p_i^{\alpha_i}}$ sont constructibles. □

Théorème 20.87 (Gauss-Wantzel[43]).
Soit $\alpha \in \mathbb{N}^*$.

(1) L'angle $\frac{2\pi}{2^\alpha}$ est constructible.

*(2) Si p est premier $p \neq 2$ 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 $\beta \in \mathbb{N}$.*

(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 de Fermat premiers 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).

Sens direct Nous supposons que l'angle $\frac{2\pi}{p^\alpha}$ est constructible ; alors le nombre $\cos\left(2\pi/p^\alpha\right)$ l'est également et le théorème de Wantzel 20.83 nous indique que

$$\left[\mathbb{Q}\Big(\cos(2\pi/p^\alpha)\Big) : \mathbb{Q}\right] = 2^m \tag{20.285}$$

pour un certain m. Posons $q = p^\alpha$ et $\omega = e^{2i\pi/q}$. Grâce au corollaire 20.24, nous savons que le polynôme minimal de ω est le polynôme cyclotomique ϕ_q dont le degré est [26]

$$\varphi(q) = \varphi(p^\alpha) = p^{\alpha-1}(p-1). \tag{20.286}$$

Par conséquent

$$\big[\mathbb{Q}(\omega) : \mathbb{Q}\big] = p^{\alpha-1}(p-1). \tag{20.287}$$

Mais par ailleurs

$$\omega + \omega^{-1} = 2\cos(2\pi/q), \tag{20.288}$$

donc $\cos(2\pi/q) \in \mathbb{Q}(\omega)$. Et en multipliant (20.288) par ω nous trouvons le polynôme annulateur suivant pour ω :

$$\omega^2 - 2\cos\left(\frac{2\pi}{q}\right)\omega + 1 = 0. \tag{20.289}$$

Cela signifie que ω est de degré 2 dans $\mathbb{Q}\big(\cos(2\pi/q)\big)$, c'est-à-dire

$$\big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/q)\big)\big] = 2. \tag{20.290}$$

En remettant bout à bout et en utilisant la propriété multiplicative des degrés des extensions [27],

$$\underbrace{\big[\mathbb{Q}(\omega) : \mathbb{Q}\big]}_{p^{\alpha-1}(p-1)} = \underbrace{\big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/p^\alpha)\big)\big]}_{2}\underbrace{\big[\mathbb{Q}\big(\cos(2\pi/p^\alpha)\big) : \mathbb{Q}\big]}_{2^m}, \tag{20.291}$$

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 :

$$p = 2^{\lambda 2^\beta} + 1 = \left(2^{2^\beta}\right)^\lambda + 1, \tag{20.292}$$

mais le lemme 3.181(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 λ soit être égal à 1 et nous avons

$$p = 1 + 2^{2^\beta}, \tag{20.293}$$

ce qui signifie que p est un nombre de Fermat premier.

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 ω est le polynôme cyclotomique ϕ_p de degré $p-1$; donc

$$[\mathbb{Q}(\omega) : \mathbb{Q}] = p - 1. \tag{20.294}$$

Un groupe d'automorphismes Nous considérons le groupe $G = \mathrm{Aut}_{\mathbb{Q}}\big(\mathbb{Q}(\omega)\big)$ 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 ω (proposition 6.85), 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 ϕ_p commute avec g, donc

$$\phi_p\big(g(\omega)\big) = g\big(\phi_p(\omega)\big) = g(0) = 0, \tag{20.295}$$

ce qui signifie que $g(\omega)$ est une racine du polynôme cyclotomique ϕ_p. Par définition 20.17, les racines sont $\{\omega, \omega^2, \ldots, \omega^{p-1}\}$, ce qui signifie que l'action de g consiste à élever ω à

26. Voir juste en dessous de la définition 20.17.
27. Proposition 6.57.

une certaine puissance entre 1 et $p-1$. Le groupe G est donc d'ordre $|G| = p-1$ et a pour éléments

$$g_k : \omega \mapsto \phi^k \tag{20.296}$$

avec $k = 1, \ldots, p-1$. L'application [28]

$$\begin{aligned} \psi : G &\to (\mathbb{Z}/p\mathbb{Z})^* \\ g_k &\mapsto [k]_p. \end{aligned} \tag{20.297}$$

est un morphisme surjectif entre deux groupes finis de même cardinal, donc c'est un isomorphisme. Le corollaire 20.34 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.

La tour d'extensions À partir de cet élément $g_0 \in G$ nous définissons avec $0 \leqslant i \leqslant n$:

$$\mathbb{K}_i = \{z \in \mathbb{Q}(\omega) \text{ tel que } g_0^{2^i}(z) = z\}. \tag{20.298}$$

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 $g_0^2(zz') = g_0^2(z)g_0^2(z')$.

$\mathbb{K}_n = \mathbb{Q}(\omega)$ 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 $z \in \mathbb{Q}(\omega)$.

$\mathbb{K}_0 = \mathbb{Q}$ Vu que 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 6.85(2) que $\{\omega^i\}_{i=0,\ldots,p-2}$ est une base [29] de $\mathbb{Q}(\omega)$. Vu que $g_0 : \mathbb{Q}(\omega) \to \mathbb{Q}(\omega)$ est un isomorphisme, l'image d'une base est une base :

$$\{g_0(\omega)^i\}_{i=0,\ldots,p-2} \tag{20.299}$$

est également une base de $\mathbb{Q}(\omega)$. Soit $z \in \mathbb{K}_0$ et décomposons-le dans cette base [30] :

$$z = \sum_{i=0}^{p-2} \lambda_i g_0(\omega)^i; \tag{20.300}$$

nous appliquons g_0 à cette égalité en tenant compte du fait que $g_0(z) = z$:

$$z = \sum_{i=0}^{p-2} \lambda_i g_0(\omega)^{i+1}. \tag{20.301}$$

En identifiant les coefficients (et en remarquant que le dernier terme dans (20.300) est le premier dans (20.301)) nous voyons que tous les coefficients sont égaux :

$$z = \lambda_0 g_0(\omega + \cdots + \omega^{p-1}) = \lambda_0 g_0\big(\phi_p(\omega) - 1\big) = -\lambda_0 \in \mathbb{Q}. \tag{20.302}$$

Dans cette chaîne d'égalités nous avons utilisé le fait que $\phi_p(X) = 1 + X + \cdots + X^{p-1}$ (corollaire 20.22) et que $\phi_p(\omega) = 0$ par définition.

$\mathbb{K}_i \subset \mathbb{K}_{i+1}$ **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 :

$$g_0^{2^{i+1}}(z) = g_0^{2\times 2^i}(z) = (g_0^{2^i})^2(z) = z. \tag{20.303}$$

Afin de voir que l'inclusion est stricte nous montrons que l'élément

$$x = \sum_{k=0}^{2^{n-i}-1} g_0^{k2^{i+1}}(\omega) \tag{20.304}$$

28. Pour rappel, la notation $[k]_p$ est la classe de l'entier k modulo p, qui est un élément de $\mathbb{Z}/p\mathbb{Z}$.
29. Ici [43] 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.
30. Il s'agit d'une base en tant qu'espace vectoriel sur $\mathbb{Q}$, donc les $\lambda ß$ sont dans $\mathbb{Q}$.

est dans $\mathbb{K}_{i+1}\backslash\mathbb{K}_i$. Dans la base $\{1,\omega,\ldots,\omega^{p-1}\}$ l'élément x a une composante ω avec le terme $k=0$, mais si on lui applique $g_0^{2^i}$ nous obtenons

$$g_0^{2^i}(x)=\sum_{k=0}^{2^{n-i-1}-1}g_0^{2^i(1+2k)}(\omega). \tag{20.305}$$

Une composante ω pour cela demanderait d'avoir $2^i(1+2k)=\lambda 2^n$ avec $\lambda\in\mathbb{N}$; cela demanderait

$$k=\lambda 2^{n-i-1}-\frac{1}{2} \tag{20.306}$$

ce qui est impossible. Donc $x\notin\mathbb{K}_i$.

Nous montrons que $x\in\mathbb{K}_{i+1}$ de la même manière :

$$g_0^{2^{i+1}}(x)=\sum_{k=0}^{2^{n-i-1}-1}g^{2^{i+1}(k+1)}(\omega). \tag{20.307}$$

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 ω dans (20.304), et il se trouve dans (20.307) avec $k=2^{n-i-1}-1$. Donc $g_0^{2^{i+1}}(x)=x$ et $x\in\mathbb{K}_{i+1}$.

Les degrés dans la tour Par définition de n et de ω, et par la propriété multiplicative des degrés[31] nous avons :

$$2^n=p-1=\big[\mathbb{Q}(\omega):\mathbb{Q}\big]=\underbrace{\big[\mathbb{Q}(\omega):\mathbb{K}_{n-1}\big]\ldots\big[\mathbb{K}_1:\mathbb{Q}\big]}_{n\text{ facteurs}}. \tag{20.308}$$

Un produit de n facteurs entiers tous strictement plus grand que zéro doit valoir 2^n. Ils doivent donc tous valoir 2 et nous avons en particulier

$$\big[\mathbb{K}_{n-1}:\mathbb{Q}\big]=2^{n-1}. \tag{20.309}$$

$\cos(2\pi/p)\in\mathbb{K}_{n-1}$ Soit $f=g_0^{2^n-1}$; vu que l'action de g_0 (comme de tous les éléments de G) est de décaler les puissances de ω, il existe $\lambda\in\mathbb{N}$ tel que $f(\omega)=\omega^\lambda$. Mais $f^2=\mathrm{Id}$ donc

$$\omega=f^2(\omega)=\omega^{2\lambda}, \tag{20.310}$$

ce qui donne $1=\omega^{\lambda^2-1}$ ou encore que p divise λ^2-1 parce que ω est une racine primitive p^{e}de l'unité. Par conséquent $[\lambda^2-1]_p=0$ ce qui donne $[\lambda]_p=\pm[1]_p$, mais comme $f\neq\mathrm{Id}$ nous avons

$$[\lambda]_p=-[1]_p. \tag{20.311}$$

Cela fait $f(\omega)=\omega^{-1}$.

Par ailleurs $\mathbb{Q}(\omega)$ est un corps, donc il contient

$$\cos\left(\frac{2\pi}{p}\right)=\frac{1}{2}(\omega+\omega^{-1}). \tag{20.312}$$

Nous en déduisons que $\cos(2\pi/p)$ est un point fixe de $f=g_0^{2^{-1}}$:

$$f\big(\cos(2\pi/p)\big)=\frac{1}{2}\big(f(\omega)+f(\omega^{-1})\big)=\cos(2\pi/p). \tag{20.313}$$

Être un point fixe de $g_0^{2^{n-1}}$ signifie être un élément de $\mathbb{K}_{n-1}$:

$$\cos\left(\frac{2\pi}{p}\right)\in\mathbb{K}_{n-1}. \tag{20.314}$$

31. Proposition 6.57

Questions de degrés Nous avons alors les (non)inclusions suivantes :

$$\mathbb{Q}\left(\cos(2\pi/p)\right) \subset \mathbb{K}_{n-1} \subsetneq \mathbb{K}_n = \mathbb{Q}(\omega), \tag{20.315}$$

ce qui fait que

$$1 < \big[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}\big] \leqslant \big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/p)\big)\big] = 2 \tag{20.316}$$

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 (20.290) (ici $\alpha = 1$). Les inégalités de (20.316) sont donc en réalité des égalités :

$$\big[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}\big] = \big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/p)\big)\big] = 2. \tag{20.317}$$

Cela, combiné au fait que $\mathbb{Q}\big(\cos(2\pi/p)\big) \subset \mathbb{K}_{n-1}$ donne

$$\mathbb{K}_{n-1} = \mathbb{Q}\big(\cos(2\pi/p)\big). \tag{20.318}$$

Mais nous savons déjà que $\big[\mathbb{K}_{n-1} : \mathbb{Q}\big] = 2^{n-1}$, donc

$$\big[\mathbb{Q}\big(\cos(2\pi/p)\big) : \mathbb{Q}\big] = 2^{n-1} \tag{20.319}$$

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 20.83.

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 20.86), 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 avec n impair côtés est constructible alors $n = \prod_i p_i^{\alpha_i}$, dont l'angle est constructible si et seulement si $\alpha = 1$ 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 $\prod_i p_i^{\alpha_i}$ est constructible : il suffit de regrouper les côtés par deux.

□

Remarque 20.88.
D'après Wikipédia[348], les seuls nombres de Fermat connus pour être premiers sont

$$F_0 = 3 \tag{20.320a}$$
$$F_1 = 5 \tag{20.320b}$$
$$F_2 = 17 \tag{20.320c}$$
$$F_3 = 257 \tag{20.320d}$$
$$F_4 = 65537. \tag{20.320e}$$

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 20.89.
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 21

Intégration sur des variétés

21.1 Variétés

21.1.1 Introduction

Soit $f : S^2 \to \mathbb{R}$ 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.

21.1.2 Définition, carte

Définition 21.1.
*Une **variété** de dimension m de classe C^1 est une partie M de $\mathbb{R}^n$ ($n \geqslant m$) munie d'un ensemble de paires $\{U_\alpha, \varphi_\alpha\}$ où U_α est un ouvert de $\mathbb{R}^n$ et $\varphi_\alpha \colon U_\alpha \to \mathbb{R}^n$ est une application vérifiant*

(1) φ_α est une bijection entre U_α et $\varphi_\alpha(U_\alpha)$;

(2) pour tout $a \in M$, il existe un α tel que $a \in \varphi_\alpha(U_\alpha)$;

(3) pour tout α, β, la partie $\varphi_\alpha^{-1}\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\big)$ est un ouvert de U_α ;

(4) L'application

$$\varphi_\beta^{-1} \circ \varphi_\alpha \colon \varphi_\alpha^{-1}\big(\varphi_\alpha(U_\alpha) \cap \varphi_\beta(U_\beta)\big) \to U_\beta \tag{21.1}$$

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 $\varphi_\alpha \colon U_\alpha \to M$ 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 φ_α peuvent transporter la mesure de U_α (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 $\mathbb{R}^n$ ».

Définition 21.2.
*Soit une variété M de dimension m. Une **carte** pour M est un couple (U, φ) où U est un ouvert de $\mathbb{R}^m$ et $\varphi\colon U \to M$ est une application*

(1) qui est une bijection entre U et son image,

(2) pour tout α,

$$\varphi_\alpha^{-1} \circ \varphi\colon \varphi^{-1}\big(\varphi(U) \cap \varphi_\alpha(U_\alpha)\big) \to \varphi_\alpha^{-1}\big(\varphi(U) \cap \varphi_\alpha(U_\alpha)\big) \tag{21.2}$$

est un C^1-difféomorphisme.

Proposition 21.3 ([1]).
Si (U_1, φ_1) et (U_2, φ_2) sont des cartes pour la variété M, alors en posant $S = \varphi_1(U_1) \cap \varphi_2(U_2)$, l'application

$$\varphi_2^{-1} \circ \varphi_1\colon \varphi_1^{-1}(S) \to \varphi_2^{-1}(S) \tag{21.3}$$

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 choisit 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 α 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

$$\varphi_2^{-1} \circ \varphi_1\colon \varphi_1\big(\varphi_\alpha(A)\big) \to \varphi_2^{-1}\big(\varphi_\alpha(A)\big) \tag{21.4}$$

est un C^1-difféomorphisme. Nous avons

$$\varphi_2 \circ \varphi_1 = \varphi_2^{-1} \circ \varphi_\alpha \circ \varphi_\alpha^{-1} \circ \varphi_1, \tag{21.5}$$

mais vu que φ_1 et φ_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. □

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 21.2 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 21.4.
Soit une variété M de dimension m et une carte $\varphi_\alpha\colon U_\alpha \to M$. Si U est un ouvert de $\mathbb{R}^p$ et si $\varphi\colon U \to M$ 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

$$\varphi_\alpha^{-1} \circ \varphi\colon \varphi^{-1}(A) \to \varphi_\alpha^{-1}(A) \tag{21.6}$$

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 12.144 implique que $m = p$. □

21.1.3 Ancienne définition

Cette section est à recycler.

Définition 21.5.
Soit $\varnothing \neq M \subset \mathbb{R}^n$, $1 \leqslant m < n$ et $k \geqslant 1$. 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 $f : V \subset \mathbb{R}^m \to \mathbb{R}^{n-m}$ 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*

$$M \cap U = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \text{ tel que } (x_{i_1}, \dots, x_{i_m}) \in V \quad \left\{ \begin{array}{rcl} x_{i_{m+1}} & = & f_1(x_{i_1}, \dots, x_{i_m}) \\ \vdots & = & \vdots \\ x_{i_n} & = & f_{n-m}(x_{i_1}, \dots, x_{i_m}) \end{array} \right. \right\}$$

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 21.6.
Soit $M \subset \mathbb{R}^n$ et $1 \leqslant m \leqslant n-1$. L'ensemble M est une variété si et seulement si $\forall a \in M$, il existe un voisinage ouvert $\mathcal{U}$ de a dans $\mathbb{R}^n$ et une application $F : W \subset \mathbb{R}^m \to \mathbb{R}^n$ où W est un ouvert tels que

(1) F est un homéomorphisme de W vers $M \cap \mathcal{U}$,

(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 $T : \mathbb{R}^m \to \mathbb{R}^n$ est une application linéaire, son rang[1] 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[2].

La condition de rang maximum sert à éviter le genre de cas de la figure 21.1 qui représente l'image de l'ouvert $]-1, 1[$ par l'application $F(t) = (t^2, t^3)$.

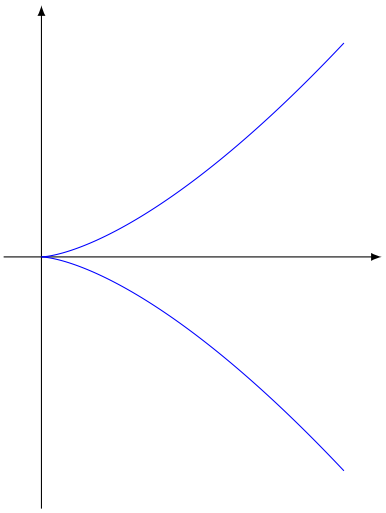


FIGURE 21.1 – Quelque chose qui n'est pas de rang maximum et qui n'est pas une variété.

La différentielle a pour matrice

$$dF(t) = (2t, 3t^2). \tag{21.7}$$

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 4.38.
2. Proposition 4.111

Proposition 21.7.
Soit $M \in \mathbb{R}^n$ et $1 \leqslant m \leqslant n-1$. 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) $M \cap \mathcal{U} = \{x \in \mathcal{U}$ tel que $G(x) = 0\}$.

21.1.4 Espace tangent

Soit M, une variété dans $\mathbb{R}^n$, et considérons un chemin $\gamma\colon I \to \mathbb{R}^n$ tel que $\gamma(t) \in M$ pour tout $t \in I$ et tel que $\gamma(0) = a$ et que γ est dérivable en 0. La **tangente** au chemin γ au point $a \in M$ est la droite

$$s \mapsto a + s\gamma'(0). \tag{21.8}$$

L'**espace tangent** de M au point a est l'ensemble décrit par toutes les tangentes en a pour tous les chemins γ possibles.

Proposition 21.8.
Une variété de dimension m dans $\mathbb{R}^n$ a un espace tangent de dimension m en chacun de ses points.

21.2 Intégration

21.2.1 Résultats préliminaires

Soient une variété M et une partie A de M. Si nous avons une fonction $f\colon A \to \mathbb{R}$, 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 φ_α nous aurions tendance à vouloir écrire

$$\int_A f = \int_{\varphi_\alpha^{-1}(A)} f \circ \varphi_\alpha \tag{21.9}$$

pour la mesure de Lebesgue induite de $\mathbb{R}^m$ vers son ouvert U_α. Pas de problème à ce que (21.9) soit bien définie. Le problème est que α n'est pas a priori spécial par rapport à β et qu'il faudrait également que

$$\int_A f = \int_{\varphi_\beta^{-1}(A)} f \circ \varphi_\beta \tag{21.10}$$

Mais en utilisant le changement de variable (théorème 15.252(3)) pour le C^1-difféomorphisme $\phi = \varphi_\alpha^{-1} \circ \varphi_\beta$, nous avons

$$\int_{\varphi_\alpha^{-1}(A)} f \circ \varphi_\alpha = \int_{\varphi_\beta(A)} (f \circ \varphi_\alpha \circ \varphi_\alpha^{-1} \circ \varphi_\beta)|J_\phi| = \int_{\varphi_\beta(A)} (f \circ \varphi_\beta)|J_\phi|. \tag{21.11}$$

À moins d'une coïncidence extraordinaire sur les valeurs du jacobien, il n'y aura pas égalité entre (21.9) et (21.10).

Pour faire mieux, il faudra ajouter quelque chose qui compense l'arrivée du jacobien. C'est ce que nous allons faire maintenant.

Proposition-définition 21.9 ([1])**.**
Soient des ouverts U_1 et U_2 de $\mathbb{R}^m$ ainsi que $A \subset \mathbb{R}^n$. Nous supposons que les applications $\varphi_1\colon U_1 \to A$ et $\varphi_2\colon U_2 \to A$ sont des bijections telles que $\varphi_1^{-1} \circ \varphi_2\colon U_2 \to U_1$ soit un C^1-difféomorphisme.

Soit une fonction mesurable $f\colon A \to [0, +\infty]$ [3].

3. Si f est seulement mesurable, allez voir les hypothèses du théorème 15.252(3).

Alors[4]

$$\int_{U_1}(f\circ\varphi_1)(x)\sqrt{\det\big((d\varphi_1)_x^*(d\varphi_1)_x\big)}dx=\int_{U_2}(f\circ\varphi_2)(y)\sqrt{\det\big((d\varphi_2)_y^*(d\varphi_2)_y\big)}dy \tag{21.12}$$

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*

$$\int_A f. \tag{21.13}$$

Démonstration. Nous considérons le C^1-difféomorphisme $\phi=\varphi_1^{-1}\circ\varphi_2$. Notre but est d'utiliser le théorème de changement de variables 15.252(2).

Vu que φ_1 est une bijection, $(\varphi_1\circ\varphi_1^{-1})(A)=A$ et nous avons

$$\phi^{-1}\big(\varphi_1^{-1}(A)\big)=\varphi_2^{-1}(A). \tag{21.14}$$

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 carré.

L'application $x\mapsto\det\big((d\varphi_1)_x^*\circ(d\varphi_1)_x\big)$ doit être composée avec ϕ pour donner

$$y\mapsto\det\big((d\varphi_1)_{\phi(y)}^*\circ(d\varphi_1)_{\phi(y)}\big). \tag{21.15}$$

Attardons-nous un peu sur $(d\varphi_1)_{\phi(y)}$. Nous savons par la règle de différentiation en chaîne 12.141 que

$$d(\varphi_1\circ\phi)_y=(d\varphi_1)_{\phi(y)}\circ(d\phi)_y. \tag{21.16}$$

Le lemme 12.143 nous dit que l'application linéaire $d\phi_y\colon\mathbb{R}^m\to\mathbb{R}^m$ est inversible et que $d\phi_y^{-1}=(d\phi^{-1})_{\phi(y)}$. En composant les deux côtés de (21.16) par cela nous trouvons

$$(d\varphi_1)_{\phi(y)}=d(\varphi_1\circ\phi)_y\circ(d\phi)_y^{-1}=(d\varphi_2)_y\circ(d\phi^{-1})_{\phi(y)}. \tag{21.17}$$

C'est me moment d'utiliser la proposition 11.74 à propos de la composition des applications adjointes. Ce qui est dans le déterminant est :

$$\big((d\varphi_2)_y\circ(d\phi^{-1})_{\phi(y)}\big)^*\circ\big((d\varphi_2)_y\circ(d\phi^{-1})_{\phi(y)}\big)=\underbrace{(d\phi^{-1})_{\phi(y)}^*}_{\mathbb{R}^m\to\mathbb{R}^m}\circ\underbrace{(d\varphi_2)_y^*\circ(d\varphi_2)_y}_{\mathbb{R}^m\to\mathbb{R}^m}\circ\underbrace{(d\phi^{-1})_{\phi(y)}}_{\mathbb{R}^m\to\mathbb{R}^m}. \tag{21.18}$$

Nous utilisons à présent un combo entre les propriétés des déterminants et de l'adjoint : les propositions 11.52(2) et 11.73 font sortir $|\det\big((d\phi^{-1})_{\phi(y)}\big)|$ de la racine carré. Ne pas oublier la valeur absolue parce que $\sqrt{x^2}=|x|$ et non $\sqrt{x^2}=x$. Cela pour dire que

$$\sqrt{\det\big((d\varphi_1)_{\phi(y)}^*\circ(d\varphi_1)_{\phi(y)}\big)}=|\det\big((d\phi^{-1})_{\phi(y)}\big)|\sqrt{\det\big((d\varphi_2)_y^*\circ(d\varphi_2)_y\big)}. \tag{21.19}$$

En ce qui concerne l'intégrale que nous voulons calculer, en mettant les bouts ensemble,

$$\begin{aligned}&\int_{\varphi_1^{-1}(A)}(f\circ\varphi_1)(x)\sqrt{\det(\cdots)}dx\\&\quad=\int_{\varphi_2^{-1}(A)}(f\circ\varphi_2)(y)|\det\big((d\phi^{-1})_{\phi(y)}\big)|\sqrt{\det\big((d\varphi_2)_y^*\circ(d\varphi_2)_y\big)}|\det(d\phi_y)|dy.\end{aligned} \tag{21.20}$$

Enfin, les déterminants se suppriment : dans la valeur absolue nous avons :

$$\det\big((d\phi^{-1})_{\phi(y)}\big)\det(d\phi_y)=\det\big((d\phi^{-1})_{\phi(y)}\circ d\phi_y\big)=\det(\mathrm{Id})=1. \tag{21.21}$$

□

4. Voir la définition de l'adjoint 11.71.

Si nous voulons que la formule de la proposition 21.9 fournisse une définition raisonnable de $\int_A f$, il faut que si (U,φ) est une carte,

$$\det\left(d\varphi_x^* \circ d\varphi_x\right) \geqslant 0 \tag{21.22}$$

pour tout x, et ne soit nul sur aucun ouvert de U.

Lemme 21.10 ([1]).
Soit une carte (U,φ) d'une variété M de dimension $\mathbb{R}^m$. 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

$$\det(d\varphi_x^* \circ d\varphi_x) > 0. \tag{21.23}$$

Démonstration. Point par point.

(1) Soit $x \in U$. Nous notons $A = d\varphi_x$. Par la proposition 11.74 nous avons $(AA^*)^* = AA^*$ du fait que $(A^*)^* = A$. Donc A^*A est autoadjoint. Sa matrice est donc symétrique par la proposition 11.71(2).
De plus, pour tout $u \in \mathbb{R}^m$ nous avons

$$\langle A^*Au, u\rangle = \langle Au, Au\rangle \geqslant 0. \tag{21.24}$$

Cela implique que la matrice de A^*A est semi-définie positive par le lemme 11.196(2). L'opérateur A^*A est également semi-défini positif par la proposition 11.195.

(2) L'existence de la base de vecteurs propres est le théorème 11.189. 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. Voila 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 un noyau (proposition 11.52(2)), c'est-à-dire un vecteur u tel que $d\varphi_a^* d\varphi_a(u) = 0$ en particulier,

$$0 = \langle d\varphi_a^* d\varphi_a u, u\rangle = \langle d\varphi_a u, d\varphi_a u,\rangle = \|d\varphi_a u\|^2. \tag{21.25}$$

Donc $d\varphi_a(u) = 0$.
Mais $\varphi\colon U \to \varphi(U)$ est bijective, donc l'application $\varphi^{-1}\colon \varphi(U) \to U$ l'est aussi et vu que φ est une carte, l'application $\varphi^{-1} \circ \varphi\colon U \to U$ est égale à l'identité. L'identité est une application linéaire dont la différentielle est elle-même, c'est-à-dire que

$$d(\varphi^{-1} \circ \varphi)_a = (d\varphi^{-1})_{\varphi(a)} \circ d\varphi_a = \mathrm{Id}\,. \tag{21.26}$$

Si donc $u \neq 0$, nous devons avoir $d\varphi_a(u) \neq 0$.

□

21.2.2 Quelque 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 de coefficient pour les variétés de petite dimension ?

C'est ce que nous allons voir maintenant.

21.2.2.1 En dimension un

Nous supposons une variété de dimension un dans $\mathbb{R}^n$. Une carte est donc une application $\varphi\colon \mathbb{R} \to \mathbb{R}^n$ dont nous notons φ_i les composantes. Nous avons, pour $u \in \mathbb{R}$,

$$d\varphi_a(u) = \frac{d}{dt}\Big[\varphi(a+tu)\Big]_{t=0} = \frac{d}{dt}\begin{pmatrix}\varphi_1(a+tu)\\ \vdots \\ \varphi_n(a+tu)\end{pmatrix} = u\begin{pmatrix}\varphi_1'(a)\\ \vdots \\ \varphi_n'(a)\end{pmatrix} = u\varphi'(a). \tag{21.27}$$

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 $x \in \mathbb{R}^m$. Pour tout $u \in \mathbb{R}$ nous avons

$$\langle d\varphi_a x, u\rangle_{\mathbb{R}} = \langle x, d\varphi_a(u)\rangle_{\mathbb{R}^n} = u\langle x, \varphi'(a)\rangle_{\mathbb{R}^n} = u\varphi'(a)\cdot x = \langle \varphi'(a)\cdot x, u\rangle_{\mathbb{R}}. \tag{21.28}$$

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 1×1 et elle se calcule en appliquant l'application au vecteur de base 1 de $\mathbb{R}$:

$$(d\varphi_a^* \circ d\varphi_a)(1) = d\varphi_a^*\big(\varphi'(a)\big) = \varphi'(a)\cdot\varphi'(a) = \|\varphi'(a)\|^2. \tag{21.29}$$

Nous avons donc

$$\det\big(d\varphi_a^* \circ d\varphi_a\big) = \|\varphi'(a)\|^2. \tag{21.30}$$

21.2.2.2 En dimension quelconque

Soit une carte $\varphi\colon \mathbb{R}^m \to \mathbb{R}^n$ pour une variété de dimension m. Les éléments de matrice de la différentielle d'une application sont donnés par la proposition 13.187 :

$$(d\varphi_a)_{ij} = \frac{\partial\varphi_i}{\partial x_j}(a). \tag{21.31}$$

Les éléments de la matrice de $d\varphi_a^*$ sont ceux de la matrice transposée. En ce qui concerna la composition, c'est le produit des matrices :

$$(d\varphi_a^* \circ d\varphi_a)_{ij} = \sum_k (d\varphi_a^*)_{ik}(d\varphi_a)_{kj} = \sum_k (d\varphi_a)_{ki}(d\varphi_a)_{kj} = \sum_k \frac{\partial\varphi_k}{\partial x_i}(a)\frac{\partial\varphi_k}{\partial x_j}(a). \tag{21.32}$$

Cela pour écrire cette bonne formule :

$$(d\varphi_a^* \circ d\varphi_a)_{ij} = \frac{\partial\varphi}{\partial x_i}(a)\cdot\frac{\partial\varphi}{\partial x_j}(a). \tag{21.33}$$

21.2.2.3 En dimension deux

Nous considérons maintenant une carte $\varphi\colon U \to \mathbb{R}^3$ avec $U \subset \mathbb{R}^2$. En notant $v_i = \partial_i\varphi(a)$ nous notons la formule (21.33) sous la forme

$$\det\big(d\varphi_a^* \circ d\varphi_a\big) = \det\begin{pmatrix} v_1\cdot v_1 & v_1\cdot v_2\\ v_2\cdot v_1 & v_2\cdot v_2\end{pmatrix} = \|v_1\|^2\|v_2\|^2 - (v_1\cdot v_2)^2. \tag{21.34}$$

L'identité de Lagrange de la proposition 11.36 nous donne alors

$$\sqrt{\det(d\varphi_a^* \circ d\varphi_a)} = \|\frac{\partial\varphi}{\partial x}\times\frac{\partial\varphi}{\partial y}\|. \tag{21.35}$$

21.2.2.4 En dimension trois

Nous considérons maintenant une carte $\varphi\colon U \to \mathbb{R}^3$ 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 11.34 à propos du produit mixte, la formule 21.33 donne

$$\det(d\varphi_a^* \circ d\varphi_a) = \det\begin{pmatrix} \|u\|^2 & u\cdot v & u\cdot w \\ v\cdot u & \|v\|^2 & v\cdot w \\ w\cdot u & w\cdot v & \|w\|^2 \end{pmatrix} = \big|(u\times v)\cdot w\big|. \tag{21.36}$$

Donc

$$\det(d\varphi_a^* \circ d\varphi_a) = \big|(\frac{\partial\varphi}{\partial x}(a)\times\frac{\partial\varphi}{\partial y}(a))\cdot\frac{\partial\varphi}{\partial z}(a)\big|. \tag{21.37}$$

21.3 Intégrale sur une variété

21.3.1 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 une vecteur normal au plan décrit par v et w qui jouit de l'importante propriété suivante :

$$\text{aire du parallélogramme} = \|v\times w\|. \tag{21.38}$$

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 $F\colon W \to F(w)$ 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

$$\frac{\partial F}{\partial x}(w), \quad \text{et} \quad \frac{\partial F}{\partial y}(w). \tag{21.39}$$

L'**élément infinitésimal de surface** sur M au point $p = F(w)$ est alors défini par

$$d\sigma_F = \|\frac{\partial F}{\partial x}(w)\times\frac{\partial F}{\partial y}(w)\|dw, \tag{21.40}$$

et si la partie $A \subset M$ est entièrement contenue dans $F(W)$, nous définissons la **mesure** de A par

$$\mu_2(A) = \int_{F^{-1}(A)} d\sigma_F = \int_{F^{-1}(A)} \|\frac{\partial F}{\partial x}(w)\times\frac{\partial F}{\partial y}(w)\|dw. \tag{21.41}$$

Remarque 21.11.
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 $\partial_x F$ et $\partial_y F$ 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 (21.41) soit indépendant du choix de F. Nous allons bientôt voir comme cas particulier du théorème 21.14 que c'est en fait le cas. C'est-à-dire que si F et $\tilde{F}$ sont deux cartes qui contiennent A, alors

$$\int_{F^{-1}(A)} d\sigma_F = \int_{\tilde{F}^{-1}(A)} d\sigma_{\tilde{F}}. \tag{21.42}$$

21.3.1.1 Exemple : la mesure de la sphère

Nous nous proposons maintenant de calculer la surface de la sphère $S^2 = x^2 + y^2 + z^2 = R^2$. L'application $F\colon B((0,0),R) \to R^3$ donnée par

$$F(x,y) = \begin{pmatrix} x \\ y \\ \sqrt{R^2 - x^2 - y^2} \end{pmatrix} \tag{21.43}$$

est une carte pour une demi-sphère. Ses dérivées partielles sont

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ -\frac{x}{\sqrt{R^2-x^2-y^2}} \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ -\frac{y}{\sqrt{R^2-x^2-y^2}} \end{pmatrix}. \tag{21.44}$$

Le produit vectoriel de ces deux vecteurs tangents donne

$$\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y) = \frac{x}{\alpha}e_1 + \frac{y}{\alpha}e_2 + e_3 \tag{21.45}$$

où $\alpha = \sqrt{R^2 - x^2 - y^2}$. En calculant la norme, nous trouvons

$$\|\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)\| = \sqrt{\frac{R^2}{R^2 - x^2 - y^2}}, \tag{21.46}$$

et en passant aux coordonnées polaires, nous écrivons l'intégrale (21.41) sous la forme

$$\int_B \|\partial_x F \times \partial_y F\| = \int_0^{2\pi} d\theta \int_0^R r\sqrt{\frac{R^2}{R^2 - x^2 - y^2}} dr = 2\pi R^2, \tag{21.47}$$

qui est bien la mesure de la demi-sphère.

21.3.2 Intégrale sur une carte

Nous pouvons maintenant définir l'intégrale d'une fonction sur une carte de la variété M.

Définition 21.12.
*Soit $F\colon W \subset \mathbb{R}^2 \to \mathbb{R}^3$, 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 $f\colon A \to \mathbb{R}$, une fonction continue. L'**intégrale** de f sur A est le nombre*

$$\int_A f = \int_A f d\sigma_F = \int_{F^{-1}(A)} (f \circ F)(w) \|\frac{\partial F}{\partial x}(w) \times \frac{\partial F}{\partial y}(w)\| dw \tag{21.48}$$

Remarque 21.13.
L'intégrale (21.48) n'est pas toujours bien définie. Étant donné que F est C^1 et que f est continue, l'intégrante 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 21.14.
Soient $F\colon W \to F(w)$ et $\tilde{F}\colon \tilde{W} \to \tilde{F}(\tilde{W})$, 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

$$\int_A f d\sigma_F = \int_A f d\sigma_{\tilde{F}}. \tag{21.49}$$

21.3.3 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

$$\begin{aligned} F\colon]0,1[\times]0,2[&\to K \\ (x,y) &\mapsto (x,y,1). \end{aligned} \tag{21.50}$$

Nous trouvons aisément les vecteurs tangents qui forment l'élément de surface :

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \tag{21.51}$$

donc $d\sigma_F = 1 \cdot dxdy$, et

$$\int_K f d\sigma_F = \int_{]0,1[\times]0,2[} f(x,y,1) \cdot 1 \cdot dxdy. \tag{21.52}$$

Nous pouvons également utiliser la carte

$$\begin{aligned} \tilde{F}\colon]0,\frac{1}{2}[\times]0,6[&\to K \\ (\tilde{x},\tilde{y}) &\mapsto (2\tilde{x},\frac{\tilde{y}}{3},1). \end{aligned} \tag{21.53}$$

Les vecteurs tangents sont maintenant

$$\frac{\partial \tilde{F}}{\partial \tilde{x}} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \frac{\partial \tilde{F}}{\partial \tilde{y}} = \begin{pmatrix} 0 \\ 1/3 \\ 0 \end{pmatrix}, \tag{21.54}$$

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

$$\int_K f d\sigma_{\tilde{F}} = \int_{]0,\frac{1}{2}[\times]0,6[} f\big(2\tilde{x},\frac{\tilde{y}}{3},1\big) \cdot \frac{2}{3} \cdot d\tilde{x}s\tilde{y}. \tag{21.55}$$

Conformément au théorème 21.14, cette dernière intégrale est égale à l'intégrale (21.52) parce qu'il s'agit juste d'un changement de variable.

21.3.4 Orientation

Soient $F\colon W \to F(w)$ et $\tilde{F}\colon \tilde{W} \to \tilde{F}(\tilde{W})$, deux cartes de la variété M. Nous pouvons considérer la fonction $h = \tilde{F}^{-1} \circ F$, définie uniquement sur l'intersection des cartes :

$$h\colon F^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big) \to \tilde{F}^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big). \tag{21.56}$$

Nous disons que F et $\tilde{F}$ ont même **orientation** si

$$J_h(w) > 0 \tag{21.57}$$

pour tout $w \in F^{-1}\big(F(W) \cap \tilde{F}(\tilde{W})\big)$.

Considérons les deux cartes suivantes pour le même carré :

$$\begin{aligned} F\colon]0,1[\times]0,1[&\to \mathbb{R}^3 \\ (x,y) &\mapsto (x,y,0) \end{aligned} \tag{21.58}$$

et

$$\begin{aligned} \tilde{F}\colon]0,\frac{1}{2}[\times]0,\frac{1}{3}[&\to \mathbb{R}^3 \\ (x,y) &\mapsto (2x,3y,0) \end{aligned} \tag{21.59}$$

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

$$\frac{\partial F}{\partial x} \times \frac{\partial F}{\partial y} = e_3, \tag{21.60}$$

tandis que

$$\frac{\partial \tilde{F}}{\partial x} \times \frac{\partial \tilde{F}}{\partial y} = 6e_3. \tag{21.61}$$

Les vecteurs normaux à la paramétrisation pointent dans le même sens.

Si par contre nous prenons la paramétrisation

$$\begin{aligned} G\colon]0,1[\times]0,1,[&\to \mathbb{R}^2 \\ (x,y) &\mapsto (x,(1-y),0), \end{aligned} \tag{21.62}$$

nous avons

$$\frac{\partial G}{\partial x} \times \frac{\partial G}{\partial y} = -e_3, \tag{21.63}$$

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.

Définition 21.15.
*Une variété M est **orientable** s'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 21.16.
Soit M, une variété orientable et un atlas orienté $\{F_i\colon W_i \to \mathbb{R}^3\}$. Alors le vecteur unitaire

$$\frac{\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)}{\left\| \frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y) \right\|} \tag{21.64}$$

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,

$$\frac{\partial (F_2 \circ h)}{\partial x} = \frac{\partial F_2}{\partial x}\frac{\partial h_1}{\partial x} + \frac{\partial F_2}{\partial y}\frac{\partial h_2}{\partial x}. \tag{21.65}$$

Après avoir fait le même calcul pour $\frac{\partial (F_2 \circ h)}{\partial y}$, nous pouvons écrire

$$\partial_x(F_2 \circ h) \times \partial_y(F_2 \circ h) = (\partial_x h_1 \partial_x F_2 + \partial_x h_2 \partial_y F_2) \times (\partial_y h_1 \partial_x F_2 + \partial_y h_2 \partial_y F_2). \tag{21.66}$$

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

$$\left(\frac{\partial F_2}{\partial x} \times \frac{\partial F_2}{\partial y}\right)(\partial_x h_1 \partial_y h_2 - \partial_x h_2 \partial_y h_2). \tag{21.67}$$

Par conséquent,

$$\frac{\partial F_1}{\partial x} \times \frac{\partial F_1}{\partial y} = \frac{\partial (F_2 \circ h)}{\partial x} \times \frac{\partial (F_2 \circ h)}{\partial y} = \left(\frac{\partial F_2}{\partial x} \times \frac{\partial F_2}{\partial y}\right) \det J_h. \tag{21.68}$$

Donc, tant que J_h est positif, les vecteurs unitaires correspondants au membre de gauche et de droite sont égaux. □

Corollaire 21.17.
Si nous avons choisi un atlas orienté pour la variété M, nous avons une fonction continue $G\colon M \to \mathbb{R}^3$ telle que $\|G(p)\| = 1$ pour tout $p \in M$. Cette fonction est donnée par

$$G(F(x,y)) = \frac{\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)}{\|\frac{\partial F}{\partial x}(x,y) \times \frac{\partial F}{\partial y}(x,y)\|} \tag{21.69}$$

sur l'image de la carte F.

Démonstration. La fonction G est construite indépendamment sur chaque carte $F(W)$ en utilisant la formule (21.69). 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. □

Il est possible que prouver, bien que cela soit plus compliqué, que la réciproque est également vraie.

Proposition 21.18.
Une variété M de dimension 2 dans $\mathbb{R}^3$ est orientable si et seulement s'il existe une fonction continue $G\colon M \to \mathbb{R}^3$ telle que pour tout $p \in M$, le vecteur $G(p)$ soit de norme 1 et normal à M au point p.

21.3.5 Formes différentielles

Nous allons donner une toute petite introduction aux formes différentielles sur des variétés compactes.

Lemme 21.19 ([349])**.**
Soit ω une k-forme sur $\mathbb{R}^n$ et f, une fonction C^∞ 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 $g\colon \mathbb{R}^n \to \mathbb{R}$. Nous avons

$$d(d^*g)X = d(g \circ f)X = (dg \circ df)X = dg\big(dfX\big) = (f^*dg)(X). \tag{21.70}$$

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

$$\omega = g\, dx^1 \wedge \dots dx^p \tag{21.71}$$

où g est une fonction C^∞. Pour soulager les notations nous allons noter $dx^I = dx^1 \wedge \dots dx^{p-1}$. Nous avons

$$\begin{aligned}
d(f^*\omega) &= d\big(f^*(gdx^I \wedge dx^p)\big) && \text{(21.72a)}\\
&= d\big(f^*(gdx^I) \wedge f^*dx^p\big) && \text{(21.72b)}\\
&= d\big(f^*(gdx^I)\big) \wedge f^*dx^p + (-1)^{p-1} f^*(gdx^I) \wedge (f^*dx^p) && \text{(21.72c)}\\
&= f^*\big(d(gdx^I)\big) \wedge f^*dx^p && \text{(21.72d)}\\
&= f^*\big(d(gdx^I) \wedge dx^p\big) && \text{(21.72e)}\\
&= f^*d\omega && \text{(21.72f)}
\end{aligned}$$

Justifications : (21.72c) est la formule de Leibnitz. (21.72d) est parce que le second terme est nul : $d(f^*dx^p) = f^*(d^2x^p) = 0$. Nous avons utilisé l'hypothèse de récurrence et le fait que $d^2 = 0$. L'étape (21.72f) est une utilisation à l'envers de la règle de Leibnitz en tenant compte que $d^2x^p = 0$. □

Lemme-définition 21.20.
Soient une variété M munie d'un recouvrement par un nombre fini d'ouverts[5] *$\{U_i\}$. Alors il existe une famille de fonctions $f_i \in C^\infty(M)$ telle que*

5. Si M n'est pas compacte, alors il faut utiliser une version un peu plus élaborée du lemme[349].

(1) $\operatorname{supp} f_i \subset U_i$,

(2) *pour tout* i, *nous avons* $f_i \geqslant 0$,

(3) *pour tout* $p \in M$ *nous avons* $\sum_i f_i(p) = 1$.

La famille (f_i) est une **partition de l'unité** subordonnée au recouvrement $\{U_i\}$.

Soit M une variété de dimension n et ω une n-forme différentielle

$$\omega_p = f(p)dx_1 \wedge \ldots \wedge dx_n. \tag{21.73}$$

Si (U, φ) est une carte ($U \subset \mathbb{R}^n$ et $\varphi\colon U \to M$) alors nous définissons

$$\int_{\varphi(U)} \omega = \int_U f\big(\varphi(x)\big)dx_1 \ldots dx_n. \tag{21.74}$$

Lorsque nous voulons intégrer sur une partie plus grande qu'une carte nous utilisons une partition de l'unité du lemme 21.20.

Définition 21.21 (Intégrale d'une forme sur une variété)**.**
Si $\{f_i\}$ *est une partition de l'unité subordonnée à un atlas de* M *nous définissons*

$$\int_M \omega = \sum_i \int_{U_i} f\omega. \tag{21.75}$$

Il est possible de montrer que cette définition ne dépend pas du choix de la partition de l'unité.

Remarque 21.22.
Nous ne définissons pas d'intégrale de k-forme différentielle sur une variété de dimension $n \neq k$. Le seul cas où cela se fait est le cas de 0-formes (les fonctions), mais cela n'est pas vraiment un cas particulier vu que les 0-formes sont associées aux n-formes de façon évidente.

Remarque 21.23.
La définition 21.21 permet d'intégrer des formes sur des variétés, pas sur des sous-variétés. Autrement dit, ce n'est pas cette définition là qu'il faut utiliser pour comprendre des objets comme

$$\int_\gamma f \tag{21.76}$$

où γ est un chemin dans $\mathbb{R}^n$. Nous définirons ce genre de choses plus bas.

21.3.6 Intégrale d'une fonction sur une sous-variété

Nous allons nous restreindre au cas d'une sous-variété de $\mathbb{R}^n$. C'est-à-dire que nous considérons l'espace $\mathbb{R}^n$ comme espace ambiant et nous allons intégrer sur des parties de $\mathbb{R}^n$. Ces parties peuvent être de dimension plus basses que n, et c'est justement ça qui fait la différence entre ce que nous faisons maintenant et la définition 21.21.

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 $F_i\colon W_i \to F_i(W_i)$.

Si $A \subset M$ est tel que pour chaque i, $A \cap F_i(W_i) = F_i(V_i)$ pour une ensemble V_i mesurable dans $\mathbb{R}^2$, alors nous considérons

$$A_1 = A \cap F_1(W_2) = F_1(V_1). \tag{21.77}$$

Ensuite, nous construisons A_2 en considérant $F_A(W_2)$ et en lui retranchant A_1 :

$$A_2 = \big(A \cap F_2(W_2)\big) \cap F_1(V_1). \tag{21.78}$$

En continuant de la sorte, nous construisons la décomposition

$$A = A_1 \cup \ldots \cup A_p \tag{21.79}$$

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 21.24.
*Si $f : A \to \mathbb{R}$ est une fonction continue, alors l'**intégrale** est le nombre*

$$\int_A f = \sum_{i=1}^{p} \int_{A_i} f d\sigma_{F_i}. \tag{21.80}$$

21.4 Longueur, aire, volumes etc.

Grâce à la mesure de Lebesgue (définition 15.210), 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 21.25.
*L'**aire** de la partie mesurable S de $\mathbb{R}^2$ est le nombre*

$$\int_{\mathbb{R}^2} \mathbb{1}_S \tag{21.81}$$

au sens de l'intégrale pour la mesure de Lebesgue de la fonction caractéristique de la partie S.

Définition 21.26.
*Le **volume** de la partie mesurable V de $\mathbb{R}^3$ est le nombre*

$$\int_{\mathbb{R}^R} \mathbb{1}_V \tag{21.82}$$

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

$$\int_{\mathbb{R}^3} \mathbb{1}_S = 0. \tag{21.83}$$

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.

21.4.1 Quelques aires faciles

Nous nous souvenons de la proposition 11.31 qui donnait une bonne propriété du produit vectoriel dans $\mathbb{R}^3$. Nous en donnons une autre.

Proposition 21.27.
Si u et v sont des vecteurs de $\mathbb{R}^2$ alors le vecteur $u \times v$ (où nous avons prolongé u et v par zéro dans $\mathbb{R}^3$) est de norme égale à l'aire[6] du parallélogramme construit sur u et v dans $\mathbb{R}^2$.

La subtilité quant au prolongement de u et v par zéro est due au fait que nous n'avons encore défini d'aires que celles de parties de $\mathbb{R}^2$. Officiellement, nous ne savons pas encore ce que serait l'aire du parallélogramme construit sur deux vecteurs de $\mathbb{R}^3$. Officieusement, nous nous attendons à ce que ce soit égal à la norme du produit vectoriel.

6. Définition 21.25.

21.5 Autres théorèmes de points fixes

En termes de théorème de points fixes nous avons déjà vu le théorème de Picard 18.34. Voir aussi le thème 18.

21.5.1 Brouwer

Proposition 21.28.
Soit $f\colon [a,b]\to[a,b]$ une fonction continue. Alors f accepte un point fixe.

Démonstration. En effet si nous considérons $g(x)=f(x)-x$ alors nous avons $g(a)=f(a)-a\geqslant 0$ et $g(b)=f(b)-b\leqslant 0$. 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 13.50. □

Exemple 21.29
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 $\cos(x)=x$. △

Proposition 21.30 (Brouwer dans $\mathbb{R}^n$ version C^∞ via Stokes)**.**
Soit B la boule fermée de centre 0 et de rayon 1 de $\mathbb{R}^n$ et $f\colon B\to B$ une fonction C^∞. Alors f admet un point fixe.

Démonstration. Supposons que f ne possède pas de points fixes. Alors pour tout $x\in B$ 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 ∂B en un point que nous appelons $g(x)$. Prouvons que cette fonction est C^k dès que f est C^k (y compris avec $k=\infty$).

Le point $g(x)$ est la solution du système

$$\begin{cases} g(x)-f(x)=\lambda\big(x-f(x)\big) & \text{(21.84a)}\\ \|g(x)\|^2=1 & \text{(21.84b)}\\ \lambda\geqslant 0. & \text{(21.84c)}\end{cases}$$

En substituant nous obtenons l'équation

$$P_x(\lambda)=\|\lambda\big(x-f(x)\big)+f(x)\|^2-1=0, \tag{21.85}$$

ou encore

$$\lambda^2\|x-f(x)\|^2+2\lambda\big(x-f(x)\big)\cdot f(x)+\|f(x)\|^2-1=0. \tag{21.86}$$

En tenant compte du fait que $\|f(x)\|<1$ (parce que les images de f sont dans $\mathcal{B}$), nous trouvons que $P_x(0)\leqslant 0$ et $P_x(1)\leqslant 0$. De même $\lim_{\lambda\to\infty}P_x(\lambda)=+\infty$. Par conséquent le polynôme de second degré P_x a exactement deux racines distinctes $\lambda_1\leqslant 0$ et $\lambda_2\geqslant 1$. La racine que nous cherchons est la seconde. Le discriminant est strictement positif, donc pas besoin d'avoir peur de la racine dans

$$\lambda(x)=\frac{-\big(x-f(x)\big)\cdot f(x)+\sqrt{\Delta_x}}{\|x-f(x)\|^2} \tag{21.87}$$

où

$$\Delta_x=4\Big(\big(x-f(x)\big)\cdot f(x)\Big)^2-4\|x-f(x)\|^2\big(\|f(x)\|^2-1\big). \tag{21.88}$$

Notons que la fonction $\lambda(x)$ est C^k dès que f est C^k ; et en particulier elle est C^∞ si f l'est.

En résumé la fonction g ainsi définie vérifie deux propriétés :

(1) elle est C^∞ ;
(2) elle est l'identité sur ∂B.

La suite de la preuve consiste à montrer qu'une telle rétraction sur B ne peut pas exister[7].

Nous considérons une forme de volume ω sur ∂B : l'intégrale de ω sur ∂B est l'aire de ∂B qui est non nulle. Nous avons alors la contradiction suivante :

$$0 < \int_{\partial B} \omega = \int_{\partial B} g^*\omega = \int_B d(g^*\omega) = \int_B g^*(d\omega) = 0 \tag{21.89}$$

Justifications :

— L'intégrale $\int_{\partial B} \omega$ est l'aire de ∂B et est donc non nulle.
— La fonction g est l'identité sur ∂B. Nous avons donc $\omega = g^*\omega$.
— Le lemme 21.19.
— La forme ω est de volume, par conséquent de degré maximum et $d\omega = 0$.

□

Un des points délicats est de se ramener au cas de fonctions C^∞. Pour la régularisation par convolution, voir [350] ; pour celle utilisant le théorème de Weierstrass, voir [351].

Théorème 21.31 (Brouwer dans $\mathbb{R}^n$ version continue)**.**
Soit B la boule fermée de centre 0 et de rayon 1 de $\mathbb{R}^n$ et $f : B \to B$ une fonction continue[8]*. Alors f admet un point fixe.*

Démonstration. Nous commençons par définir une suite de fonctions

$$f_k(x) = \frac{f(x)}{1 + \frac{1}{k}}. \tag{21.90}$$

Nous avons $\|f_k - f\|_\infty \leqslant \frac{1}{1+k}$ où la norme est la norme uniforme sur B. Par le théorème de Weierstrass 13.304 il existe une suite de fonctions $C^\infty(B, \mathbb{R})$ que nous nommons g_k telles que

$$\|g_k - f_k\|_\infty \leqslant \frac{1}{1+k}. \tag{21.91}$$

Vérifions que cette fonction g_k soit bien une fonction qui prend ses valeurs dans B :

$$\begin{aligned}
\|g_k(x)\| &\leqslant \|g_k(x) - f_k(x)\| + \|f_k(x)\| && \text{(21.92a)}\\
&\leqslant \frac{1}{1+k} + \frac{\|f(x)\|}{1 + \frac{1}{k}} && \text{(21.92b)}\\
&\leqslant \frac{1}{1+k} + \frac{1}{1 + \frac{1}{k}} && \text{(21.92c)}\\
&= 1. && \text{(21.92d)}
\end{aligned}$$

Par la version C^∞ du théorème (proposition 21.30), g_k admet un point fixe que l'on nomme x_k.

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 :

$$\begin{aligned}
\|f(x) - x\| &= \|f(x) - g_k(x) + g_k(x) - x_k + x_k - x\| && \text{(21.93a)}\\
&\leqslant \|f(x) - g_k(x)\| + \underbrace{\|g_k(x) - x_k\|}_{=0} + \|x_k - x\| && \text{(21.93b)}\\
&\leqslant \frac{1}{1+k} + \|x_k - x\|. && \text{(21.93c)}
\end{aligned}$$

En prenant le limite $k \to \infty$ le membre de droite tend vers zéro et nous obtenons $f(x) = x$. □

7. 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.
8. Une fonction continue sur un fermé de $\mathbb{R}^n$ est à comprendre pour la topologie induite.

21.5.2 Théorème de Schauder

Une conséquence du théorème de Brouwer est le théorème de Schauder qui est valide en dimension infinie.

Théorème 21.32 (Théorème de Schauder[352])**.**
Soit E, un espace vectoriel normé, K un convexe compact de E et $f\colon K \to K$ une fonction continue. Alors f admet un point fixe.

Démonstration. Étant donné que $f\colon K \to K$ est continue, elle y est uniformément continue. Si nous choisissons ϵ alors il existe $\delta > 0$ tel que

$$\|f(x) - f(y)\| \leqslant \epsilon \tag{21.94}$$

dès que $\|x - y\| \leqslant \delta$. La compacité de K permet de choisir un recouvrement fini par des ouverts de la forme

$$K \subset \bigcup_{1\leqslant i\leqslant p} B(x_j, \delta) \tag{21.95}$$

où $\{x_1, \ldots, x_p\} \subset K$. Nous considérons maintenant $L = \text{Span}\{f(x_j) \text{ tel que } 1 \leqslant j \leqslant p\}$ et

$$K^* = K \cap L. \tag{21.96}$$

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 7.59). 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 (21.95) sur K (voir le lemme 21.20). Nous commençons par définir

$$\psi_j(x) = \begin{cases} 0 & \text{si } \|x - x_j\| \geqslant \delta \\ 1 - \frac{\|x-x_j\|}{\delta} & \text{sinon.} \end{cases} \tag{21.97}$$

pour chaque $1 \leqslant j \leqslant p$. Notons que ψ_j est une fonction positive, nulle en-dehors de $B(x_j, \delta)$. En particulier la fonction suivante est bien définie :

$$\varphi_j(x) = \frac{\psi_j(x)}{\sum_{k=1}^p \psi_k(x)} \tag{21.98}$$

et nous avons $\sum_{j=1}^p \varphi_j(x) = 1$. Les fonctions φ_j sont continues sur K et nous définissons finalement

$$g(x) = \sum_{j=1}^p \varphi_j(x) f(x_j). \tag{21.99}$$

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 10.26.

Nous considérons seulement la restriction $g\colon K^* \to K^*$ qui est continue sur un compact contenu dans un espace vectoriel de dimension finie. Le théorème de Brouwer nous enseigne alors que g a un point fixe (proposition 21.31). Nous nommons y ce point fixe. Notons que y est fonction du ϵ choisi au début de la construction, via le δ qui avait conditionné la partition de l'unité.

Nous avons

$$f(y) - y = f(y) - g(y) \tag{21.100a}$$

$$= \sum_{j=1}^p \varphi_j(y) f(y) - \sum_{j=1}^p \varphi_j(y) f(x_j) \tag{21.100b}$$

$$= \sum_{j=1}^p \varphi(j)(y)\big(f(y) - f(x_j)\big). \tag{21.100c}$$

Par construction, $\varphi_j(y) \neq 0$ seulement si $\|y - x_j\| \leqslant \delta$ et par conséquent seulement si $\|f(y) - f(x_j)\| \leqslant \epsilon$. D'autre par nous avons $\varphi_j(y) \geqslant 0$; en prenant la norme de (21.100) nous trouvons

$$\|f(y) - y\| \leqslant \sum_{j=1}^{p} \|\varphi_j(y)\big(f(y) - f(x_j)\big)\| \leqslant \sum_{j=1}^{p} \varphi_j(y)\epsilon = \epsilon. \tag{21.101}$$

Nous nous souvenons maintenant que y était fonction de ϵ. Soit y_m le y qui correspond à $\epsilon = 2^{-m}$. Nous avons alors

$$\|f(y_m) - y_m\| \leqslant 2^{-m}. \tag{21.102}$$

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 $y^* \in K$ [9]. Nous avons les majorations

$$\|f(y^*) - y^*\| \leqslant \|f(y^*) - f(y_m)\| + \|f(y_m) - y_m\| + \|y_m - y^*\|. \tag{21.103}$$

Si m est assez grand, les trois termes du membre de droite peuvent être rendus arbitrairement petits, d'où nous concluons que

$$f(y^*) = y^* \tag{21.104}$$

et donc que f possède un point fixe. □

21.5.3 Théorème de Cauchy-Arzella

Théorème 21.33 (Cauchy-Arzela[282]).
Nous considérons le système d'équation différentielles

$$\begin{cases} y' = f(t, y) & \text{(21.105a)} \\ y(t_0) = y_0. & \text{(21.105b)} \end{cases}$$

avec $f : U \to \mathbb{R}^n$, 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 (21.105) existe.

Idée de la démonstration. Nous considérons $M = \|f\|_\infty$ et K, l'ensemble des fonctions M-Lipschitz sur U. Nous prouvons que $(K, \|.\|_\infty)$ est compact. Ensuite nous considérons l'application

$$\begin{aligned} \Phi &: K \to K \\ \Phi(f)(t) &= x_0 + \int_{t_0}^{t} f\big(u, f(u)\big)du. \end{aligned} \tag{21.106}$$

Après avoir prouvé que Φ était continue, nous concluons qu'elle a un point fixe par le théorème de Schauder 21.32. □

Remarque 21.34.
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 33.15 et 33.16.

9. Notons que même dans la sous-suite nous avons $\|f(y_m) - y_m\| \leqslant 2^{-m}$, avec le même « m » des deux côtés de l'inégalité.

21.5.4 Théorème de Markov-Kakutani et mesure de Haar

Définition 21.35.
*Soit G un groupe topologique. Une **mesure de Haar** sur G est une mesure positive*[10] *μ telle que*

(1) $\mu(gA) = \mu(A)$ pour tout mesurable A et tout $g \in G$,

(2) $\mu(K) < \infty$ pour tout compact $K \subset G$.

Si de plus le groupe G lui-même est compact nous demandons que la mesure soit normalisée : $\mu(G) = 1$.

Le théorème suivant nous donne l'existence d'une mesure de Haar sur un groupe compact.

Théorème 21.36 (Markov-Katutani[353])**.**
Soit E un espace vectoriel normé et L, une partie non vide, convexe, fermée et bornée de E'. Soit $T : L \to L$ une application continue. Alors T a un point fixe.

Démonstration. Nous considérons un point $x_0 \in L$ et la suite

$$x_n = \frac{1}{n+1} \sum_{i=0}^{n} T^i x_0. \tag{21.107}$$

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

$$C = \bigcap_{n \in \mathbb{N}} \overline{\{x_m \text{ tel que } m \geqslant n\}}. \tag{21.108}$$

Le lemme 9.53 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

$$\lim_{n \to \infty} x_{\sigma(n)}(v) = x(v) \tag{21.109}$$

pour tout $v \in E$. Montrons que x est un point fixe de T. Nous avons

$$\|(Tx_{\sigma(k)} - x_{\sigma(k)})v\| = \Big\| T \frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^i x_0(v) - \frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^i x_0(v) \Big\| \tag{21.110a}$$

$$= \Big\| \frac{1}{1+\sigma(k)} \sum_{i=0}^{\sigma(k)} T^{i+1} x_0(v) - T^i x_0(v) \Big\| \tag{21.110b}$$

$$= \frac{1}{1+\sigma(k)} \big\| T^{\sigma(k)+1} x_0(v) - x_0(v) \big\| \tag{21.110c}$$

$$\leqslant \frac{2M}{\sigma(k)+1} \tag{21.110d}$$

où $M = \sum_{y \in L} \|y(v)\| < \infty$ parce que L est borné. En prenant $k \to \infty$ nous trouvons

$$\lim_{k \to \infty} \big(Tx_{\sigma(k)} - x_{\sigma(k)} \big) v = 0, \tag{21.111}$$

ce qui signifie que $Tx = x$ parce que T est continue. □

Le théorème suivant est une conséquence du théorème de Markov-Katutani.

Théorème 21.37.
Si G est un groupe topologique compact possédant une base dénombrable de topologie alors G accepte une unique mesure de Haar normalisée. De plus elle est unimodulaire :

$$\mu(Ag) = \mu(gA) = \mu(A) \tag{21.112}$$

pour tout mesurables $A \subset G$ et tout élément $g \in G$.

10. Définition 15.20.

21.6 Intégrales curvilignes

21.6.1 Chemins de classe C^1

Définition 21.38.
*Soit $p, q \in \mathbb{R}^n$. Un **chemin** C^1 par morceaux joignant p à q est une application continue*

$$\gamma : [a,b] \to \mathbb{R}^n \qquad \gamma(a) = p, \gamma(b) = q \tag{21.113}$$

pour laquelle il existe une subdivision $a = t_0 < t_1 < \ldots < t_{r-1} < t_r = b$ telle que :

(1) la restriction de γ sur chaque ouvert $]t_i, t_{i+1}[$ est de classe C^1 ;

(2) pour tout $0 \leqslant i \leqslant r$, γ' possède une limite à gauche (sauf pour $i = 0$) et une limite à droite (sauf pour $i = r$) en t_i.

*Le **chemin γ est (globalement) de classe C^1** si la subdivision peut être choisie de « longueur » $r = 1$.*

Remarque 21.39.
Si a et b sont des points de $\mathbb{R}^n$, on peut créer le chemin particulier

$$\gamma : [0,1] \to \mathbb{R}^n : t \mapsto (1-t)a + tb \tag{21.114}$$

qui relie ces points par un segment de droite.

21.6.2 Intégrer une fonction

Définition 21.40.
*Soit $f : D \subset \mathbb{R}^n \to \mathbb{R}$ une fonction continue, et $\gamma : [a,b] \to D$ un chemin C^1. On définit **l'intégrale de f sur γ** par*

$$\int_\gamma f ds = \int_\gamma f = \int_a^b f(\gamma(t)) \left\|\gamma'(t)\right\| dt. \tag{21.115}$$

Exemple 21.41
Soit l'hélice

$$\begin{aligned} \sigma\colon [0, 2\pi] &\to \mathbb{R}^3 \\ t &\mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \\ t \end{pmatrix}, \end{aligned} \tag{21.116}$$

et la fonction $f(x,y,z) = x^2 + y^2 + z^2$. L'intégrale de f sur σ est

$$\begin{aligned} \int_\sigma f &= \int_0^{2\pi} (\cos^2 t + \sin^2 t + t^2) \|\sigma'(t)\| dt \\ &= \int_0^{2\pi} (1 + t^2)\sqrt{2} dt \\ &= \sqrt{2}\left[t + \frac{t^3}{3}\right]_0^{2\pi} \\ &= \sqrt{2}\left(2\pi + \frac{8\pi^3}{8}\right). \end{aligned} \tag{21.117}$$

△

Remarque 21.42.
Si $f = 1$, alors nous tombons sur

$$\int_\gamma ds = \int_a^b \|\gamma'(t)\| dt, \tag{21.118}$$

Nous verrons par le théorème 22.10 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 15.157 en gardant en tête que $\int_\gamma 1$ n'est pas la mesure de l'image de γ dans $\mathbb{R}^2$.

Proposition 21.43 (Indépendence en la paramétrisation)**.**
La valeur de l'intégrale de f sur γ ne dépend pas du paramétrage (équivalent ou pas) choisi.

Démonstration. Soit donc un chemin $\gamma\colon [c,d] \to \mathbb{R}^3$ ainsi que $\varphi\colon [c,d] \to [a,b]$, une reparamétrisation de classe C^1, strictement monotone et le chemin σ définit par $\gamma(s) = \sigma\big(\varphi(s)\big)$ avec $s \in [c,d]$. En supposant que $\varphi'(s) \geqslant 0$, nous avons

$$\begin{aligned} I = \int_\gamma f &= \int_c^d f\big(\gamma(s)\big)\|\gamma'(s)\| ds \\ &= \int_c^d f\Big(\sigma\big(\varphi(s)\big)\Big)\|\sigma'\big(\varphi(s)\big)\||\varphi'(s)| ds. \end{aligned} \tag{21.119}$$

Pour cette intégrale, nous posons $t = \varphi(s)$, et par conséquent $dt = \varphi'(s)ds$. Étant donné que $\varphi'(s) \geqslant 0$, nous pouvons supprimer les valeurs absolues, et obtenir

$$\begin{aligned} I &= \int_{\varphi(c)}^{\varphi(d)} f\big(\sigma(t)\big)\|\sigma'(t)\| dt \\ &= \int_a^b f\big(\sigma(t)\big)\|\sigma'(t)\| dt \\ &= \int_\sigma f. \end{aligned} \tag{21.120}$$

Essayez de faire le cas $\varphi'(s) \leqslant 0$. □

Remarque 21.44.
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 γ est un chemin dans $\mathbb{C}$, l'intégrale

$$\int_\gamma f \tag{21.121}$$

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 $\gamma\colon [a,b] \to \mathbb{C}$ est un chemin, alors

$$\int_\gamma f = \int_\gamma f(z)dz = \int_a^b f\big(\gamma(t)\big)\gamma'(t)dt, \tag{21.122}$$

sans valeur absolue autour de $\gamma'(t)$.

21.6.3 Intégrer un champ de vecteurs

Définition 21.45.
*Un **champ de vecteur** est une application $G : \mathbb{R}^n \to \mathbb{R}^n$. On définit l'intégrale de G sur un chemin $\gamma : [a,b] \to \mathbb{R}^n$ par*

$$\int_\gamma G \stackrel{def}{=} \int_a^b \big\langle G(\gamma(t)), \gamma'(t) \big\rangle dt.$$

Remarque 21.46.
Cette définition ne dépend pas de la paramétrisation choisie, mais le signe change selon le sens du chemin.

Si $\sigma'(t) \neq 0$, nous pouvons considérer le vecteur unitaire tangent à la courbe :

$$T(t) = \frac{\sigma'(t)}{\|\sigma'(t)\|}. \tag{21.123}$$

Si F est un champ de vecteurs sur $\mathbb{R}^3$, la circulation de F le long de σ sera donnée par

$$\int_\sigma F \cdot ds = \int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt = \int_a^b F\big(\sigma(t)\big) \cdot \frac{\sigma'(t)}{\|\sigma'(t)\|}dt = \int_\sigma F \cdot T ds \tag{21.124}$$

où dans la dernière expression, $F \cdot T$ est vu comme fonction $(x, y, z) \mapsto F(x, y, z) \cdot T(x, y, z)$. L'intégrale d'un champ de vecteurs sur une courbe n'est donc rien d'autre que l'intégrale de la composante tangentielle du champ de vecteurs.

21.6.4 Intégrer une forme différentielle sur un chemin

La formule d'intégration d'un champ de vecteur [11],

$$\int_\gamma G = \int_{[a,b]} \langle G(\gamma(t)), \gamma'(t)\rangle dt, \tag{21.125}$$

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 13.430)

$$\begin{aligned} G_x^\flat \colon \mathbb{R}^n &\to \mathbb{R} \\ v &\mapsto \langle G(x), v\rangle \end{aligned} \tag{21.126}$$

pour écrire de façon plus compacte :

$$\int_\gamma G = \int_{[a,b]} G^\flat_{\gamma(t)}\big(\gamma'(t)\big)dt. \tag{21.127}$$

Définition 21.47.
*Soient une forme différentielle ω sur $\mathbb{R}^n$ et un chemin de classe C^1 $\gamma\colon [a, b] \to \mathbb{R}^n$. L'**intégrale** de ω sur γ est définie par*

$$\int_\gamma \omega = \int_a^b \omega_{\gamma(t)}\gamma'(t)dt \tag{21.128}$$

Remarque 21.48.
Cette définition ne dépend pas de la paramétrisation choisie, mais le signe change selon le sens du chemin.

21.6.5 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 $\gamma\colon [a, b] \to \mathbb{R}^n$ est le chemin, les formules sont

$$\begin{aligned} \int_\gamma f &= \int_{[a,b]} f\big(\gamma(t)\big)\|\gamma'(t)\|dt \\ \int_\gamma G &= \int_{[a,b]} \langle G\big(\gamma(t)\big), \gamma'(t)\rangle dt. \end{aligned} \tag{21.129}$$

11. Définition 21.45.

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 21.49.
*Soit $\gamma\colon [a,b] \to \mathbb{R}^n$, un chemin de classe C^1 tel que son image est contenue dans le domaine D. Si ω es une 1-forme différentielle sur D, nous définissons l'**intégrale de ω le long de** γ le nombre*

$$\begin{aligned}\int_\gamma \omega &= \int_a^b \omega_{\gamma(t)}\big(\gamma'(t)\big)dt \\ &= \int_a^b \Big[a_1\big(\gamma(t)\big)\gamma_1'(t) + \cdots + a_n\big(\gamma(t)\big)\gamma_n'(t)\Big]dt.\end{aligned} \tag{21.130}$$

Cette définition est une bonne définition parce que si on change la paramétrisation du chemin, on ne change pas la valeur de l'intégrale, c'est la proposition suivante.

Proposition 21.50.
Si γ et β sont des chemins équivalents, alors

$$\int_\gamma \omega = \int_\beta \omega, \tag{21.131}$$

c'est-à-dire que l'intégrale est invariante sous les reparamétrisation du chemin.

Démonstration. Deux chemins sont équivalents quand il existe un difféomorphisme C^1 $h\colon [a,b] \to [c,d]$ tel que $\gamma(t) = (\beta \circ h)(t)$. En remplaçant γ par $(\beta \circ h)$ dans la définition de $\int_\gamma \omega$, nous trouvons

$$\int_a^b \omega_{\gamma(t)}\big(\gamma'(t)\big)dt = \int_a^b \omega_{(\beta\circ h)(t)}\big((\beta\circ h)'(t)\big)dt. \tag{21.132}$$

Un changement de variable $u = h(t)$ transforme cette dernière intégrale en $\int_\beta \omega$, ce qui prouve la proposition. □

Remarque 21.51.
Si γ est une somme de chemins, $\gamma = \gamma^{(1)} + \cdots + \gamma^{(n)}$, où chacun des $\gamma^{(i)}$ est un chemin, alors

$$\int_\gamma \omega = \sum_{i=1}^n \int_{\gamma_i} \omega \tag{21.133}$$

parce que ω est linéaire.

Remarque 21.52.
Si $-\gamma$ est le chemin

$$\begin{aligned}-\gamma\colon [a,b] &\to \mathbb{R}^n \\ t &\mapsto \gamma\big(b-(t-a)\big),\end{aligned} \tag{21.134}$$

alors

$$\int_{-\gamma} \omega = -\int_\gamma \omega, \tag{21.135}$$

c'est-à-dire que si l'on parcours 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,

$$\int_\gamma G^\flat = \int_\gamma G. \tag{21.136}$$

En effet,

$$\begin{aligned}\int_{\gamma G^\flat} &= \int_a^b G^\flat_{\gamma(t)}(\gamma'(t)) \\ &= \int_a^b \big[G_1(\gamma(t))dx_1 + \dots G_n(\gamma(t))dx_n\big]\big(\gamma_1'(t), \dots, \gamma_n'(t)\big) \\ &= \int_a^b \langle G(\gamma(t)), \gamma'(t)\rangle \\ &= \int_\gamma G.\end{aligned} \tag{21.137}$$

Proposition 21.53.
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 γ.

Démonstration. Nous avons

$$\begin{aligned}\int_\gamma \omega = \int_\gamma df &= \int_a^b \sum_{i=1} n \frac{\partial f}{\partial x_i}\big(\gamma(t)\big)\gamma_i'(t)dt \\ &= \int_a^b \frac{d}{dt}\Big((f\circ\gamma)(t)\Big)dt \\ &= (f\circ\gamma)(b) - (f\circ\gamma(a)).\end{aligned} \tag{21.138}$$

□

21.6.6 Interprétation physique : travail

Définition 21.54.
Une force $F\colon D \subset \mathbb{R}^n \to \mathbb{R}^n$ est ***conservative*** *si elle dérive d'un potentiel, c'est-à-dire s'il existe une fonction $V \in C^1(D, \mathbb{R})$ telle que*

$$F(x) = (\nabla V)(x). \tag{21.139}$$

Étant donné que F est un champ de vecteurs, nous avons une forme différentielle associée $F^\flat$,

$$F^\flat_x\colon x \mapsto \langle F(x), v\rangle. \tag{21.140}$$

Lemme 21.55.
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 $F^\flat_x = dV(x)$. En effet,

$$\begin{aligned}F^\flat_x(v) &= \langle F(x), v\rangle \\ &= F_1(x)v_1 + \dots + F_n(x)v_n \\ &= \frac{\partial V}{\partial x_1}(x)v_1 + \dots \frac{\partial V}{\partial x_n}(x)v_n \\ &= dV(x)v.\end{aligned} \tag{21.141}$$

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, $F = \nabla V$. □

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 21.56.
Si F est une force conservative, alors le travail de F lors d'un déplacement ne dépend pas du chemin suivit.

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 :

$$W_\gamma(F) = \int_\gamma F = \int_\gamma dV = V\big(\gamma(b)\big) - V\big(\gamma(a)\big). \tag{21.142}$$

Donc si β est un autre chemin tel que $\beta(a) = \gamma(a)$ et $\beta(b) = \gamma(b)$, nous avons $W_\beta(F) = W_\gamma(F)$. □

21.6.7 Intégrer un champs de vecteurs sur un bord en $2D$

Si $D \subset \mathbb{R}^2$ est tel que ∂D est une variété de dimension 1 et tel que D accepte un champ de vecteur normal extérieur unitaire ν. Si nous voulons définir

$$\int_{\partial D} G, \tag{21.143}$$

le mieux est de prendre une paramétrisation $\gamma\colon [0,1] \to \mathbb{R}^2$ et de calculer

$$\int_0^1 \langle G_{\gamma(t)}, \frac{\dot\gamma(t)}{\|\dot\gamma(t)\|}\rangle dt. \tag{21.144}$$

Hélas, cette définition ne fonctionne pas parce que son signe dépend du sens de la paramétrisation γ. Si la paramétrisation tourne dans l'autre sens, il y a un signe de différence.

Nous allons définir

$$\int_{\partial D} G = \int_0^1 \langle G_{\gamma(t)}, T(t)\rangle dt \tag{21.145}$$

où $T(t) = \dot\gamma(t)/\|\dot\gamma(t)\|$ et où γ est choisi de telle façon que la rotation d'angle $\frac{\pi}{2}$ amène ν 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.

21.6.8 Intégrer une forme différentielle sur un bord en $2D$

Nous n'allons pas chercher très loin :

$$\int_{\partial D} \omega = \int_{\partial D} \omega^\sharp, \tag{21.146}$$

c'est-à-dire que l'intégrale de la forme différentielle est celle du champ de vecteur associé. Le membre de droite est définit par (21.145), avec le choix d'orientation qui va avec.

21.6.9 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 $F\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$ est la carte. Nous supposons que D vérifie les hypothèses de la formule de Green. Alors nous définissons

$$\int_{F(\partial D)} \omega = \int_{\partial D} F^*\omega \tag{21.147}$$

où $F^*\omega$ est la forme différentielle définie sur ∂D par $(F^*\omega)(v) = \omega\big(dF(v)\big)$.

Cette définition est très abstraite, mais nous n'allons, en pratique, jamais l'utiliser, grâce au théorème de Stokes.

21.6.10 Intégrer d'un champ de vecteurs sur un bord en $3D$

Encore une fois, nous n'allons pas chercher bien loin :

$$\int_{F(\partial D)} G = \int_{F(\partial D)} G^\flat \tag{21.148}$$

où $G^\flat$ est la forme différentielle associée au champ de vecteur. Le membre de droite est définit par l'équation (21.147).

21.6.11 Dérivées croisées et forme différentielle exacte

Nous considérons le problème suivant : trouver une fonction $f\colon \mathbb{R}^2 \to \mathbb{R}$ telle que

$$\begin{cases} \frac{\partial f}{\partial x} = a(x,y) & (21.149a) \\ \frac{\partial f}{\partial y} = b(x,y) & (21.149b) \end{cases}$$

où a et b sont des fonctions supposées suffisamment régulières. Nous savons que ce problème n'a pas de solutions lorsque

$$\frac{\partial a}{\partial y} \neq \frac{\partial b}{\partial x} \tag{21.150}$$

parce que cela impliquerait $\partial^2_{xy} f \neq \partial^2_{yx} f$. Nous sommes en droit de nous demander si la condition

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x} \tag{21.151}$$

impliquerait qu'il existe une solution au problème (21.149). La réponse est oui, et nous allons brièvement la justifier. Pour plus de détails nous vous demandons de chercher un peu sur internet les mots-clefs *forme différentielles exacte*. Vous consulterez également avec profit [354].

Proposition 21.57.
Si a et b sont des fonctions qui satisfont à la condition

$$\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}, \tag{21.152}$$

alors la fonction

$$f(x,y) = \int_0^x a(t,0)dt + \int_0^y b(x,t)dt \tag{21.153}$$

répond au problème

$$\begin{cases} \frac{\partial f}{\partial x} = a(x,y) & (21.154a) \\ \frac{\partial f}{\partial y} = b(x,y) & (21.154b) \end{cases}$$

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 :

$$\int_0^x \frac{\partial f}{\partial x}(t,y)dt = f(x,y) \tag{21.155}$$

et son pendant par rapport à y :

$$\int_0^y \frac{\partial f}{\partial y}(x,t)dt = f(x,y). \tag{21.156}$$

En appliquant ces version du théorème fondamental, nous obtenons immédiatement.

$$\frac{\partial f}{\partial y} = b(x,y). \tag{21.157}$$

En ce qui concerne la dérivée par rapport à y,

$$\begin{aligned} \frac{\partial f}{\partial x} &= a(x,0) + \int_0^y \frac{\partial b}{\partial x}(x,t)dt & (21.158a) \\ &= a(x,0) + \int_0^y \frac{\partial a}{\partial y}(x,t)dt & (21.158b) \\ &= a(x,0) + [a(x,t)]_{t=0}^{t=y} & (21.158c) \\ &= a(x,y). & (21.158d) \end{aligned}$$

□

En ce qui concerne l'unicité, supposons que f et g soient deux solutions au problème. L'équation

$$\frac{\partial f}{\partial x} = a(x,y) = \frac{\partial g}{\partial x} \tag{21.159}$$

implique que

$$f(x,y) = g(x,y) + C(y) \tag{21.160}$$

où C est une fonction seulement de y. L'autre équation implique

$$f(x,y) = g(x,y) + D(x) \tag{21.161}$$

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 (21.153) 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

$$F(x,y) = \begin{pmatrix} a(x,y) \\ b(x,y) \end{pmatrix}. \tag{21.162}$$

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

$$\nabla f = \begin{pmatrix} a(x,y) \\ b(x,y) \end{pmatrix}. \tag{21.163}$$

21.7 Surfaces paramétrées

De la même façon qu'un chemin dans $\mathbb{R}^3$ est décrit comme une application $\sigma\colon \mathbb{R} \to \mathbb{R}^3$, une surface dans $\mathbb{R}^3$ sera vue comme une application $\varphi\colon \mathbb{R}^2 \to \mathbb{R}^3$. Une **surface paramétrée** dans $\mathbb{R}^3$ est une application

$$\begin{aligned} \varphi\colon D \subset \mathbb{R}^2 &\to \mathbb{R}^3 \\ (u,v) &\mapsto \varphi(u,v) = \begin{pmatrix} x(u,v) \\ y(u,v) \\ z(u,z) \end{pmatrix}. \end{aligned} \tag{21.164}$$

Nous allons parler de la « surface φ » pour désigner l'image de φ dans $\mathbb{R}^3$.

Si on fixe le paramètre $u - u_0$, alors l'application

$$v \mapsto \varphi(u_0, v) \tag{21.165}$$

est un chemin dans la surface. Un vecteur tangent à ce chemin sera tangent à la courbe :

$$\frac{\partial \varphi}{\partial v}(u_0, v_0) = \begin{pmatrix} \frac{\partial x}{\partial v}(u_0, v_O) \\ \frac{\partial y}{\partial v}(u_0, v_O) \\ \frac{\partial z}{\partial v}(u_0, v_O) \end{pmatrix}. \tag{21.166}$$

De même, en fixant v_0, on considère le chemin

$$u \mapsto \varphi(u, v_0). \tag{21.167}$$

Le vecteur tangent à ce chemin est égalent tangent à la surface :

$$\frac{\partial \varphi}{\partial u} = \begin{pmatrix} \frac{\partial x}{\partial u}(u_0, v_O) \\ \frac{\partial y}{\partial u}(u_0, v_O) \\ \frac{\partial z}{\partial u}(u_0, v_O) \end{pmatrix} \tag{21.168}$$

Définition 21.58.
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 à la paramétrisation 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 :

$$n(u_0,v_0)=\frac{\partial\varphi}{\partial u}(u_0,v_0)\times\frac{\partial\varphi}{\partial v}(u_0,v_0). \tag{21.169}$$

L'équation du plan tangent est alors obtenue par

$$\begin{pmatrix}x-x_0\\y-y_0\\z-z_0\end{pmatrix}\cdot n(u_0,v_0)=0 \tag{21.170}$$

où $x_0=x(u_0,v_0)$, $y_0=y(u_0,v_0)$, $z_0=z(u_0,v_0)$.

21.7.1 Graphe d'une fonction

Soit la fonction $f\colon D\subset\mathbb{R}^2\to\mathbb{R}$. Le graphe de f est l'ensemble des points de la forme

$$\big(x,y,f(x,y)\big) \tag{21.171}$$

tels que $(x,y)\in D$. Cela est une surface paramétrée par

$$\begin{aligned}\varphi\colon D&\to\mathbb{R}^3\\(x,y)&\mapsto\begin{pmatrix}x\\y\\f(x,y)\end{pmatrix}.\end{aligned} \tag{21.172}$$

Les vecteurs tangents sont

$$\frac{\partial\varphi}{\partial x}=\begin{pmatrix}1\\0\\\frac{\partial f}{\partial x}\end{pmatrix},\quad\frac{\partial\varphi}{\partial y}=\begin{pmatrix}0\\1\\\frac{\partial\varphi}{\partial y}\end{pmatrix}. \tag{21.173}$$

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

$$n=\begin{vmatrix}e_x&e_y&e_z\\1&0&\partial_xf(x_0,y_0)\\0&1&\partial_yf(x_0,y_0)\end{vmatrix}=-\frac{\partial f}{\partial x}(x_0,y_0)e_x-\frac{\partial f}{\partial y}(x_0,y_0)e_y+e_z. \tag{21.174}$$

En suivant l'équation (21.170), nous avons l'équation suivante pour le plan :

$$\begin{pmatrix}x-x_0\\y-y_0\\z-z_0\end{pmatrix}\cdot\begin{pmatrix}-\frac{\partial f}{\partial x}(x_0,y_0)\\-\frac{\partial f}{\partial y}(x_0,y_0)\\1\end{pmatrix}=0, \tag{21.175}$$

c'est-à-dire

$$-(x-x_0)\frac{\partial f}{\partial x}(x_0,y_0)-(y-y_0)\frac{\partial f}{\partial y}(x_0,y_0)+z-f(x_0,y_0)=0, \tag{21.176}$$

ce qui revient à

$$z-f(x_0,y_0)=\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0). \tag{21.177}$$

Nous retrouvons donc l'équation du plan tangent à un graphe.

Exemple 21.59
La sphère de rayon R peut être paramétrée par les angles sphériques :

$$\phi(\theta,\varphi) = \begin{pmatrix} R\sin\theta\cos\varphi \\ R\sin\theta\sin\varphi \\ R\cos\theta \end{pmatrix} \tag{21.178}$$

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

$$T_\theta = \frac{\partial\phi}{\partial\theta}(R,\frac{\pi}{2},0) = \begin{pmatrix} R\cos\theta\cos\varphi \\ R\cos\theta\sin\varphi \\ -R\sin\theta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -R \end{pmatrix}, \tag{21.179}$$

et

$$T_\varphi = \frac{\partial\phi}{\partial\varphi}(R,\frac{\pi}{2},0) = \begin{pmatrix} -R\sin\theta\sin\varphi \\ R\sin\theta\cos\varphi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ R \\ 0 \end{pmatrix}. \tag{21.180}$$

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

$$\begin{vmatrix} e_x & e_y & e_z \\ 0 & 0 & -R \\ 0 & R & 0 \end{vmatrix} = R^2 e_x. \tag{21.181}$$

Ici encore, nous avons le vecteur que nous attendions sur un dessin. L'équation du plan tangent est maintenant

$$\begin{pmatrix} x-R \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} R^2 \\ 0 \\ 0 \end{pmatrix} = 0, \tag{21.182}$$

c'est-à-dire $R^2(x-R)=0$ et donc $x=R$. △

21.7.2 Intégrale sur une partie de $\mathbb{R}^m$

Soit M une variété de dimension n dans $\mathbb{R}^m$. Soit $F : W \subset \mathbb{R}^n \to M$ une paramétrisation 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

$$d\sigma = \sqrt{\det(J_F(w)^t J_F(w))}dw \tag{21.183}$$

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.

21.8 Intégrales de surface

21.8.1 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 apparait également dans l'intégration sur des surfaces de l'espace.

Définition 21.60.
Une ***orientation*** *sur une surface $S \subset \mathbb{R}^3$ est le choix d'un champ de vecteurs continu $\nu : S \to \mathbb{R}^3$ 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 ν 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 ν. L'intégrale de G sur S, aussi appelée le *flux de G à travers S*, est

$$\iint_S G \cdot dS \stackrel{\text{def}}{=} \iint_S \langle G, \nu \rangle \, d\sigma. \tag{21.184}$$

Si on suppose que la surface est paramétrée par une application

$$F : W \subset \mathbb{R}^2 \to \mathbb{R}^3 : (u, v) \mapsto (F_1(u, v), F_2(u, v), F_3(u, v))$$

alors un vecteur unitaire ν 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 cette paramétrisation l'intégrale (21.184) 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).

21.9 Intégrales de surface

21.9.1 Aire d'une surface paramétrée

Lorsque nous avions vu la longueur d'une courbe paramétrée, nous avions 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

$$T_u = \frac{\partial \varphi}{\partial u}(u_0, v_0) \quad T_v = \frac{\partial \varphi}{\partial v}(u_0, v_0). \tag{21.185}$$

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 :

$$dS = \|T_u \times T_v\|. \tag{21.186}$$

L'aire de la surface donné par $\varphi\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$ sera donc donnée par

$$Aire\big(\varphi(D)\big) = \iint_D \|T_u \times T_v\| du\, dv. \tag{21.187}$$

Exemple 21.61
Calculons l'aire de la sphère. Les vecteurs tangents ont déjà été calculés aux équations (21.179) et (21.180) :

$$T_\theta = \begin{pmatrix} R\cos\theta\cos\varphi \\ R\cos\theta\sin\varphi \\ -R\sin\theta \end{pmatrix}, \quad T_\varphi = \begin{pmatrix} -R\sin\theta\sin\varphi \\ R\sin\theta\cos\varphi \\ 0 \end{pmatrix}. \tag{21.188}$$

Le produit vectoriel vaut

$$\begin{aligned} T_\theta \times T_\varphi &= \begin{vmatrix} e_x & e_y & e_z \\ R\cos\theta\cos\varphi & R\cos\theta\sin\varphi & -R\sin\theta \\ -R\sin\theta\sin\varphi & R\sin\theta\cos\varphi & 0 \end{vmatrix} \\ &= (R^2\sin^2\theta\cos\varphi)e_x + (R^2\sin^2\theta\sin\varphi)e_y \\ &\quad + (R^2\cos\theta\sin\theta\cos^2\varphi + R^2\sin\theta\cos\theta\sin^2\varphi)e_z. \end{aligned} \tag{21.189}$$

La norme demande quelques calculs et mises en évidences. Le résultat est :

$$\|T_\theta \times T_\varphi\| = R^2\sin\theta. \tag{21.190}$$

L'aire de la sphère est donc donnée par

$$Aire = \int_0^{2\pi} d\varphi \int_0^\pi R^2\sin\theta d\theta = 2\pi R^2[-\cos\theta]_0^\pi = 4\pi R^2. \tag{21.191}$$

Il est bon de se souvenir que, en coordonnées sphériques,

$$\|T_\theta \times T_\varphi\| = R^2\sin\theta. \tag{21.192}$$

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

$$T_\theta \times T_\varphi = R^2\sin(\theta)e_r. \tag{21.193}$$

△

Remarque 21.62.
L'équation (21.190) 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 21.63
Nous pouvons donner l'aire du graphe d'une fonction quelconque. La surface est paramétrée par

$$\varphi(x,y) = \begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix}. \tag{21.194}$$

Les vecteurs tangents sont

$$T_x = \begin{pmatrix} 1 \\ 0 \\ \partial_x f \end{pmatrix}, \quad T_y = \begin{pmatrix} 0 \\ 1 \\ \partial_y f \end{pmatrix}. \tag{21.195}$$

Le produit vectoriel est donné par

$$T_x \times T_y = \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & \partial_x f \\ 0 & 1 & \partial_y f \end{vmatrix} = (-\partial_x f)e_x - (\partial_y f)e_y + e_z. \tag{21.196}$$

L'élément de surface est par conséquent

$$dS = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}, \tag{21.197}$$

et la surface du graphe sera

$$Aire = \iint_D \sqrt{\left(\frac{\partial f}{\partial x}(x,y)\right)^2 + \left(\frac{\partial f}{\partial y}(x,y)\right)^2 + 1}\, dx\, dy \tag{21.198}$$

△

21.9.2 Intégrale d'une fonction sur une surface

Si S est une surface dans $\mathbb{R}^3$ paramétrée par

$$\begin{aligned} &\varphi\colon D \to \mathbb{R}^3 \\ &(u,v) \mapsto \varphi(u,v) \in S, \end{aligned} \tag{21.199}$$

et si f est une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$ définie au moins sur S, l'intégrale de f sur S est logiquement définie par

$$\int_S f\, dS = \iint_D f\big(\varphi(u,v)\big) \|T_u(u,v) \times T_v(u,v)\| du dv \tag{21.200}$$

où $T_u = \frac{\partial \varphi}{\partial u}$ et $Y_v = \frac{\partial \varphi}{\partial v}$. La quantité

$$\|T_u(u,v) \times T_v(u,v)\| du dv \tag{21.201}$$

est appelé **élément de surface**.

Encore une fois, si on prend $f = 1$, alors on retrouve la surface de S :

$$\int_S dS = Aire(S). \tag{21.202}$$

Remarque 21.64.
Le nombre $\int_S f dS$ ne dépend pas de la paramétrisation choisie pour S.

21.9.3 Aire d'une surface de révolution

Soit γ une courbe dans le plan xy, paramétrée par

$$\gamma(u) = \begin{pmatrix} x(u) \\ y(u) \\ 0 \end{pmatrix} \tag{21.203}$$

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

$$\varphi(u,\theta) = \begin{pmatrix} x(u) \\ y(u)\cos\theta \\ y(u)\sin\theta \end{pmatrix} \tag{21.204}$$

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 cette paramétrisation sont

$$T_u = \frac{\partial\varphi}{\partial u} = \begin{pmatrix} x'(u) \\ y'(u)\cos\theta \\ y'(u)\sin\theta \end{pmatrix} \quad T_\theta = \frac{\partial\varphi}{\partial\theta} = \begin{pmatrix} 0 \\ -y(u)\sin\theta \\ y(u)\cos\theta \end{pmatrix}. \tag{21.205}$$

Le produit vectoriel de ces deux vecteurs vaut

$$\begin{aligned} T_u \times T_\theta &= \begin{vmatrix} e_x & e_y & e_z \\ x' & y'\cos\theta & y'\sin\theta \\ 0 & -y\sin\theta & y\cos\theta \end{vmatrix} \\ &= y'(u)y(u)\,e_x - x'(u)y(u)\cos\theta\,e_y + x'(u)y(u)\sin\theta\,e_z. \end{aligned} \tag{21.206}$$

En ce qui concerne la norme :

$$dS = \|T_u \times T_\theta\| = \sqrt{(y'y)^2 + (x'y)^2} = |y(u)|\sqrt{y'(u)^2 + x'(u)^2}. \tag{21.207}$$

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 :

$$\begin{aligned} Aire(S) &= \int_0^{2\pi} d\theta \int_a^b y(u)\sqrt{x'(u)^2 + y'(u)^2}du \\ &= 2\pi \int_a^b y(u)\sqrt{x'(u)^2 + y'(u)^2}du. \end{aligned} \tag{21.208}$$

Exemple 21.65

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 21.2.

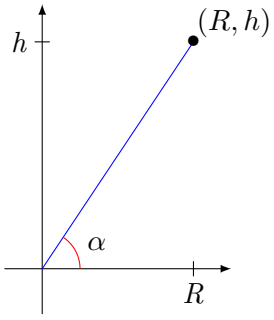


FIGURE 21.2 – En faisant tourner cette droite autour de l'axe X, nous obtenons un cône.

Ce segment peut être paramétré par

$$\gamma(u) = \begin{pmatrix} Ru \\ hu \\ 0 \end{pmatrix} \tag{21.209}$$

avec $u \in [0,1]$. Cela donne $x(u) = Ru$, $y(u) = hu$ et par conséquent

$$Aire = 2\pi \int_0^1 hu\sqrt{R^2+h^2} = \pi h\sqrt{R^2+h^2}. \tag{21.210}$$

Ce résultat peut aussi être exprimé en fonction de l'angle. En sachant que $h = \sqrt{h^2+R^2}\sin(\alpha)$, nous trouvons

$$Aire = \pi(R^2+h^2)\sin(\alpha). \tag{21.211}$$

△

Exemple 21.66
Calculons la surface latérale du tore obtenu par révolution du cercle de la figure 21.3.

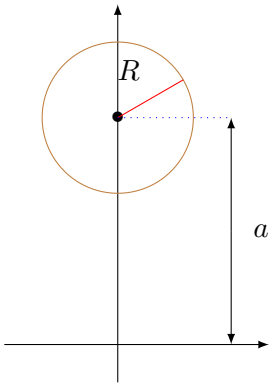


FIGURE 21.3 – Si nous tournons ce cercle autour de l'axe X, nous obtenons un tore de rayon « externe » a et de rayon « interne » R.

Le chemin qui détermine le cercle de départ est

$$\gamma(u) = \begin{pmatrix} R\cos(u) \\ a + R\sin(u) \\ 0 \end{pmatrix}, \tag{21.212}$$

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

$$\begin{aligned} Aire &= 2\pi \int_0^{2\pi} \big(a + R\sin(u)\big) R\, du \\ &= 2\pi R\big(2\pi a + R[-\cos(u)]_0^{2\pi}\big) \\ &= 4\pi^2 aR. \end{aligned} \tag{21.213}$$

△

21.9.4 Intégrale d'une 2-forme

Nous considérons ω, une 2-forme différentielle sur $\mathbb{R}^2$.

Définition 21.67.
Si $\omega_{(x,y)} = u(x,y)dx \wedge dy$ et si D est un ouvert de $\mathbb{R}^2$ alors nous définissons

$$\int_D \omega = \int_D u(x,y)dx\, dy. \tag{21.214}$$

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 $\phi\colon D \to S$ de S avec D ouvert dans $\mathbb{R}^2$ compatible avec l'orientation, c'est-à-dire que pour tout $(t,s) \in D$,

$$n\big(\phi(t,s)\big) = \frac{\partial\phi}{\partial t}(t,s) \times \frac{\partial\phi}{\partial s}(t,s). \tag{21.215}$$

Définition 21.68.
Pour intégrer ω sur S nous faisons

$$\int_S \omega = \int_D \phi^*\omega \tag{21.216}$$

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

$$(\phi^*\omega)(u,v) = \omega\big(d\phi(u), d\phi(v)\big) \tag{21.217a}$$

$$= \omega\left(u_1\frac{\partial\phi}{\partial x_1} + u_2\frac{\partial\infty}{\partial x_2}, v_1\frac{\partial\phi}{\partial x_1} + v_2\frac{\partial\infty}{\partial x_2}\right). \tag{21.217b}$$

Les termes en u_1v_1 et u_2v_2 sont nuls ; par exemple :

$$\omega\left(u_1\frac{\partial\phi}{\partial x_1}, v_1\frac{\partial\phi}{\partial x_1}\right) = u_1v_1\omega\left(\frac{\partial\infty}{\partial x_1}, \frac{\partial\phi}{\partial x_1}\right) = 0 \tag{21.218}$$

parce que ω est antisymétrique. Il nous reste donc

$$(\phi^*\omega)(u,v) = (u_1v_2 - u_2v_1)\omega\left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}\right) \tag{21.219a}$$

$$= \omega\left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_2}\right)(dt \wedge ds)(u,v). \tag{21.219b}$$

Cette dernière ligne est bien de la forme $\phi^*\omega = F(t,s)dt \wedge ds$.

21.10 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 :

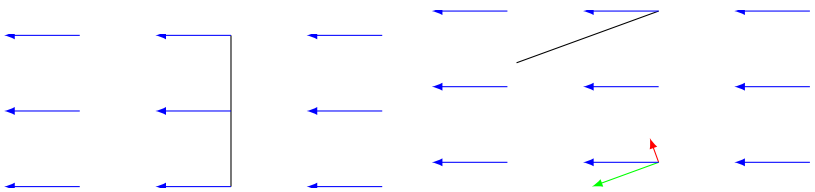
(1) plus il y a d'eau, plus ça pousse ;
(2) plus la surface de la palle est grande, plus on va utiliser d'eau ;
(3) plus la palle est perpendiculaire au courant, plus on va en profiter.

Nous voyons sur la figure 21.4 que lorsque la palle 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.

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 $\varphi\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$, 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 :

$$n(u,v) = \frac{T_u \times T_v}{\|T_u \times T_v\|}. \tag{21.220}$$



(a) L'eau pousse bien perpendiculairement à la palle du moulin.

(b) Lorsque l'eau ne pousse pas perpendiculairement, une partie de la force est perdue. Ici la partie normale (en rouge) est très petite.

FIGURE 21.4 – La partie rouge de la force est perdue si l'eau ne pousse pas perpendiculairement. De plus lorsque la palle est inclinée, elle prend moins d'eau sur elle.

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

$$F(p)_\perp = F(p) \cdot n(p). \tag{21.221}$$

C'est ce nombre là que nous intégrons sur la surface.

Définition 21.69.
*Le **flux du champ de vecteurs** à travers la surface $S = \varphi(D)$ est*

$$\int F \cdot dS = \int F \cdot n \, dudv. \tag{21.222}$$

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

$$\int F \cdot dS = \iint_D F\big(\varphi(u,v)\big) \cdot (T_u \times T_v) \, dudv \tag{21.223}$$

où $T_u = \frac{\partial \phi}{\partial u}$ et $T_v = \frac{\partial \varphi}{\partial v}$.

Exemple 21.70
Soit le champ de vecteurs

$$F = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}. \tag{21.224}$$

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 (21.193) :

$$T_\theta \times T_\varphi = R^2 \sin(\theta). \tag{21.225}$$

Nous savons aussi que

$$F\big(\phi(\theta,\varphi)\big) = 2e_r. \tag{21.226}$$

L'intégrale à calculer est donc

$$I = \int_0^\pi d\theta \int_0^{2\pi} d\varphi \, 2e_r \cdot \big(R^2 \sin(\theta) e_r\big). \tag{21.227}$$

Vu que le produit scalaire $e_r \cdot e_r$ vaut 1, nous calculons

$$I = 4\pi R^2 \int_0^\pi \sin(\theta) d\theta = 8\pi R^2. \tag{21.228}$$

△

Exemple 21.71
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

$$G(r,\theta,\varphi) = \frac{1}{r^2}e_r. \tag{21.229}$$

Sur la sphère de rayon R, nous avons

$$G\big(\phi(\theta,\varphi)\big) = \frac{1}{R^2}e_r. \tag{21.230}$$

L'intégrale est donc

$$\int_0^{\pi} d\theta \int_0^{2\pi} \frac{1}{R^2}e_r \cdot \big(R^2\sin(\theta)e_r\big)d\varphi = 8\pi. \tag{21.231}$$

Ce flux ne dépend pas de R. △

Exemple 21.72
Soit S le disque de rayon 5 placé horizontalement à la hauteur 12. Calculer le flux du champ de vecteurs

$$F(x,y,z) = xe_x + ye_y + ze_z. \tag{21.232}$$

Les équations de la surface sont $z = 12$, $x^2+y^2 \leqslant 25$. Nous prenons le paramétrage en coordonnées cylindriques :

$$\varphi(r,\theta) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ 12 \end{pmatrix}. \tag{21.233}$$

Les vecteurs tangents sont

$$T_r = \frac{\partial\varphi}{\partial r} = \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} \quad T_\theta = \frac{\partial\varphi}{\partial\theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \\ 0 \end{pmatrix}. \tag{21.234}$$

Le vecteur normal est alors

$$T_r \times T_\theta = re_z. \tag{21.235}$$

Sur la surface, le champ de vecteurs s'écrit

$$F\big(\varphi(r,\theta)\big) = r\cos(\theta)e_x + r\sin(\theta)e_y + 12e_z. \tag{21.236}$$

Par conséquent

$$F \cdot (T_r \times T_\theta) = 12r. \tag{21.237}$$

L'intégrale à calculer est

$$\begin{aligned} \int_0^5 dr \int_0^{2\pi} 12r\,d\theta &= 12 \cdot 2\pi \int_0^5 r\,dr \\ &= \frac{25}{2}24\pi \\ &= 300\pi. \end{aligned} \tag{21.238}$$

△

21.11 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 :

Théorème 21.73.
Si M est une variété orientable de dimension n avec un bord noté ∂M, alors pour toute forme différentielle ω de degré $n-1$ on a

$$\int_M d\omega = \int_{\partial M} \omega.$$

où $d\omega$ désigne la différentielle extérieure de ω.

Nous allons maintenant voir quelques cas particuliers.

Une des nombreuses formes du théorème de Stokes (théorème 21.73) est que si la forme différentielle ω est exacte alors son intégrale est facile.

Théorème 21.74 ([1])**.**
Si γ est une chemin de classe C^1 dans un ouvert Ω et si ω est la forme différentielle exacte $\omega = df$, alors

$$\int_\gamma df = f\big(\gamma(1)\big) - f\big(\gamma(0)\big). \tag{21.239}$$

Démonstration. C'est une application du lemme 13.222 et du théorème fondamental du calcul intégral 15.233. □

21.11.1 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 ∂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 ∂D, il faudra faire en sorte de ne considérer que la composante de G normale à ∂D : pas question d'intégrer par exemple la norme de G sur ∂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 21.75.
*Soit $x \in \partial D$ et $\nu \in \mathbb{R}^n$, nous disons que ν est un **vecteur normal extérieur** de ∂D si*

(1) $\langle \nu, v\rangle = 0$ pour tout vecteur tangent v à ∂D au point x. Pour rappel, ∂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 $c + t\nu \notin \bar{D}$ et $x - t\nu \in D$.

Nous pouvons maintenant définir le concept de flux. Soit $D \subset \mathbb{R}^n$ tel que ∂D soit une variété de dimension $n-1$ qui admette un vecteur normal extérieur $\nu(x)$ en chaque point. Soit aussi $G\colon \bar{D} \to \mathbb{R}^n$, un champ de vecteur de classe C^1. Le **flux** de G au travers de ∂D est le nombre

$$\int_{\partial D} \langle G(x), \nu(x)\rangle d\sigma(x). \tag{21.240}$$

Cette intégrale est en général très compliquée à calculer parce qu'il faut trouver le champ de vecteur normal, puis une paramétrisation de la surface ∂D et ensuite appliquer la méthode décrite au point 21.8.

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 21.76 (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 :

$$\partial D = A_1 \cup \dots A_p \cup N \tag{21.241}$$

où

(1) $A_1, \dots, A_p, N$ sont deux à deux disjoints,

(2) pour tout $i \leqslant p$, 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 ν.

Si G est un champ de vecteurs de classe C^1 sur $\bar{D}$ alors

$$\int_D \nabla \cdot G = \sum_{i=1}^{p} \int_{A_i} \langle G, \nu \rangle. \tag{21.242}$$

L'intégrale du membre de gauche est l'intégrale sur un ouvert d'une simple fonction.

21.11.2 Formule de Green

Pour rappel, une chemin $\gamma\colon [0,1] \to \mathbb{R}^n$ est **régulier** s'il est C^1 et si $\gamma(t) \neq 0$ pour tout r. Le chemin est de **Jordan** si $\gamma(1) = \gamma(0)$ et si $\gamma\colon [a,b[\to \mathbb{R}^n$ est injective.

La formule de Green est un cas particulier du théorème de la divergence dans le cas $n = 2$, légèrement reformulé :

Théorème 21.77.
Soit $D \subset \mathbb{R}^2$ ouvert borné tel que son bord est est la réunion finie d'un certain nombre de chemins de classe C^1 de Jordan réguliers. Supposons qu'en chaque point de son bord, D possède un vecteur normal unitaire extérieur ν. Soient P et Q deux fonctions réelles de classe C^1 sur $\bar{D}$. Alors

$$\iint_D (\partial_x Q - \partial_y P) dx\, dy = \oint_{\partial D} P dx + Q dy \tag{21.243}$$

où chaque chemin γ 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 21.73. Nous considérons la forme différentielle

$$\omega = P dx + Q dy, \tag{21.244}$$

et sa différentielle

$$d\omega = \sum_i d\omega_i \wedge dx_i \tag{21.245a}$$

$$= \left(\frac{\partial P}{\partial x} dx + \frac{\partial P}{\partial y} dy\right) \wedge dx + \left(\frac{\partial Q}{\partial x} dx + \frac{\partial Q}{\partial y} dy\right) \wedge dy \tag{21.245b}$$

$$= \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx \wedge dy. \tag{21.245c}$$

Intégrons cette forme $d\omega$ sur le domaine ouvert D que nous paramétrons de façon triviale par

$$\begin{aligned}\varphi\colon D &\to \mathbb{R}^2\\ (u,v) &\mapsto (u,v).\end{aligned} \tag{21.246}$$

Ce que nous avons est

$$\iint_D d\omega = \iint_D d\omega_{(u,v)}\left(\frac{\partial\varphi}{\partial u},\frac{\partial\varphi}{\partial v}\right)dudv \tag{21.247}$$

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

$$(dx\wedge dy)(T_u,T_v) = dx(T_u)dy(T_v) - dx(T_v)dy(T_u) = 1 - 0 = 1. \tag{21.248}$$

L'intégrale (21.247) se développe donc en

$$\iint_D d\omega = \iint_D \left(\frac{\partial Q}{\partial x}(u,v) - \frac{\partial P}{\partial y}(u,v)\right)(dx\wedge dy)(T_u,T_v)dudv = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dudv. \tag{21.249}$$

Par conséquent la formule de Stokes nous donne la formule (21.243).

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

$$\omega = -\frac{y}{2}dx + \frac{x}{2}dy, \tag{21.250}$$

de telle sorte que $\partial_x Q - \partial_y Q = 1$ et que

$$\iint_D d\omega = \iint_D ddudv = S, \tag{21.251}$$

et au final l'aire est donnée par

$$S = \int_{\partial D}\left(-\frac{y}{2}dx + \frac{x}{2}dy\right). \tag{21.252}$$

Lorsque le bord de D est paramétré par

$$\begin{aligned}\gamma\colon [a,b] &\to \mathbb{R}^2\\ u &\mapsto \begin{pmatrix}x(u)\\y(u)\end{pmatrix},\end{aligned} \tag{21.253}$$

nous avons

$$(Pdx + Qdy)\gamma'(u) = Px' + Qy', \tag{21.254}$$

et alors

$$\int_{\partial D} Pdx + Qdy = \int_a^b P\big(x(u),y(u)\big)x'(u) + Q\big(x(u),y(u)\big)y'(u)du. \tag{21.255}$$

En ce qui concerne l'aire de la surface, nous prenons les P et Q de la forme 21.250 :

$$S = \frac{1}{2}\int_a^b \Big(-y(u)x'(u) + x(u)y'(u)\Big)du. \tag{21.256}$$

21.11.3 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 quel surface dont le bord est la courbe. La version présentée ici suppose que la surface peut se paramétrer en un seul morceau :

Théorème 21.78 (Formule de Stokes)**.**
Soit $F : W \subset \mathbb{R}^2 \to \mathbb{R}^3$ une paramétrisation (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 la paramétrisation

$$N = \frac{\frac{\partial F}{\partial u} \wedge \frac{\partial F}{\partial v}}{\left\| \frac{\partial F}{\partial u} \wedge \frac{\partial F}{\partial v} \right\|} \tag{21.257}$$

alors

$$\iint_{F(D)} \langle \nabla \times G, N \rangle \, d\sigma_F = \int_{F(\partial D)} G \tag{21.258}$$

où les chemins formant le bord ∂D sont orientés comme dans le théorème de Green.

Notons, juste pour avoir une bonne nouvelle de temps en temps, que

$$d\sigma_F = \left\| \frac{\partial F}{\partial u} \times \frac{\partial F}{\partial v} \right\| du dv, \tag{21.259}$$

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[12], la formule (21.258) s'écrit

$$\int_{F(D)} \langle \nabla \times G, N(x) \rangle \, d\sigma_F(x) = \int_{F(\gamma)} \langle G, T \rangle \, ds \tag{21.260}$$

où T est le vecteur unitaire tangent à $F(\gamma)$.

21.11.3.1 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

$$F(r, \theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 1 \end{pmatrix}. \tag{21.261}$$

Mais nous aurions aussi pu mettre les coordonnées r et θ dans l'autre ordre :

$$\tilde{F}(\theta, r) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 1 \end{pmatrix}. \tag{21.262}$$

Les vecteurs normaux ne sont pas les même : 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_θ et 1_r soient dans la même orientation que les vecteurs ν et T tels que donnés par le théorème de Green, et tels que dessinés sur la figure 21.5.

Plus généralement, nous choisissons l'ordre des coordonnées u et v pour que la base $(1_u, 1_v)$ ait la même orientation que (ν, T) où T a le sens convenu dans le théorème de Green.

21.12 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.

12. et surtout plus explicite.

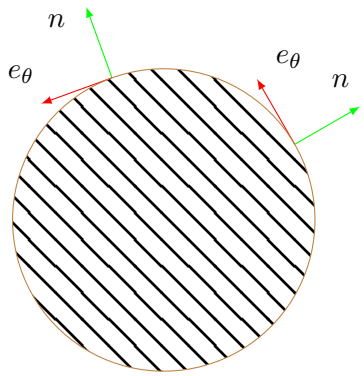


FIGURE 21.5 – L'orientation sur le cercle. Si nous les prenons dans l'ordre, les vecteurs $(1_r, 1_\theta)$ ont la même orientation que celle donnée par les vecteurs (ν, T) donnés par la convention de Green.

21.12.1 L'intégrale d'une fonction sur les réels

Si $f\colon [a,b] \subset \mathbb{R} \to \mathbb{R}$ est une fonction usuelle, sont intégrale est

$$\int_a^b f(x)dx = F(b) - F(a) \tag{21.263}$$

où F est une primitive de f.

21.12.2 Intégrale d'une fonction sur un chemin

Si f est une fonction sur $\mathbb{R}^3$ et si $\sigma\colon [a,b] \to \mathbb{R}^3$ est un chemin dans $\mathbb{R}^3$, l'intégrale de f sur σ est, par définition,

$$\int f\, d\sigma = \int_a^b f\big(\sigma(t)\big)\|\sigma'(t)\|dt. \tag{21.264}$$

21.12.3 Intégrale d'une fonction sur une surface

Nous devons paramétrer la surface S par une application $\varphi\colon D \subset \mathbb{R}^2 \to \mathbb{R}^3$. À partir d'une telle paramétrisation, nous construisons un élément de surface en prenant le produit vectoriel des deux vecteurs tangents :

$$dS = \frac{\partial\varphi}{\partial u} \times \frac{\partial\varphi}{\partial v} du dv. \tag{21.265}$$

L'intégrale est

$$\int f\, dS = \iint_D f\big(\varphi(u,v)\big) \left\| \frac{\partial\varphi}{\partial u} \times \frac{\partial\varphi}{\partial v} \right\| du dv. \tag{21.266}$$

Il ne faut pas rajouter de jacobien : la norme du produit vectoriel *est* le jacobien.

Remarque 21.79.
La formule (21.266) 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 21.80
Les coordonnées polaires sont données par

$$\varphi(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{pmatrix}. \tag{21.267}$$

Les vecteurs tangents à cette paramétrisation sont

$$T_r = \frac{\partial\varphi}{\partial r} = \begin{pmatrix}\cos\theta\\ \sin\theta\\ 0\end{pmatrix}, \quad T_\theta = \frac{\partial\varphi}{\partial\theta} = \begin{pmatrix}-r\sin\theta\\ r\cos\theta\\ 0\end{pmatrix}. \tag{21.268}$$

Le vecteur normal est

$$\frac{\partial\varphi}{\partial r} \times \frac{\partial\varphi}{\partial\theta} = \begin{vmatrix} e_x & e_y & e_z\\ \cos\theta & \sin\theta & 0\\ -r\sin\theta & r\cos\theta & 0\end{vmatrix} = re_z. \tag{21.269}$$

Nous trouvons donc que l'élément de surface est la norme de re_z, c'est-à-dire r, le jacobien connu. △

21.12.4 Intégrale d'une fonction sur un volume

Si V est un volume dans $\mathbb{R}^3$, nous effectuons la même procédure : nous trouvons une paramétrisation, et nous formons un élément de volume avec les vecteurs tangents de la paramétrisation. Nous avons donc un volume déterminé par l'application

$$\varphi\colon D \subset \mathbb{R}^3 \to \mathbb{R}^3, \tag{21.270}$$

et ses trois vecteurs tangents

$$\begin{aligned} T_u &= \frac{\partial\varphi}{\partial u}\\ T_v &= \frac{\partial\varphi}{\partial v}\\ T_w &= \frac{\partial\varphi}{\partial w}. \end{aligned} \tag{21.271}$$

Comment former un volume avec trois vecteurs ? Réponse : le produit mixte. L'intégrale de f sur V sera

$$\int f\,dV = \iiint_D f\big(\varphi(u,v,w)\big)\left\|\frac{\partial\varphi}{\partial u}\cdot\left(\frac{\partial\varphi}{\partial v}\times\frac{\partial\varphi}{\partial w}\right)\right\| dudv. \tag{21.272}$$

Encore une fois, le produit mixte *est* le jacobien. Prenons les coordonnées sphériques :

$$\begin{aligned} x(r,\theta,\varphi) &= r\sin(\theta)\cos(\varphi)\\ y(r,\theta,\varphi) &= r\sin(\theta)\sin(\varphi)\\ z(r,\theta,\varphi) &= r\cos(\theta) \end{aligned} \tag{21.273}$$

Les trois vecteurs tangents seront

$$T_r = \begin{pmatrix}\frac{\partial x}{\partial r}\\ \frac{\partial y}{\partial r}\\ \frac{\partial z}{\partial r}\end{pmatrix} = \begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix} \tag{21.274a}$$

$$T_\theta = \begin{pmatrix}\frac{\partial x}{\partial\theta}\\ \frac{\partial y}{\partial\theta}\\ \frac{\partial z}{\partial\theta}\end{pmatrix} = \begin{pmatrix}r\cos(\theta)\cos(\varphi)\\ r\cos(\theta)\sin(\varphi)\\ -r\sin(\theta)\end{pmatrix} \tag{21.274b}$$

$$T_\varphi = \begin{pmatrix}\frac{\partial x}{\partial\varphi}\\ \frac{\partial y}{\partial\varphi}\\ \frac{\partial z}{\partial\varphi}\end{pmatrix} = \begin{pmatrix}-r\sin(\theta)\sin(\varphi)\\ r\sin(\theta)\cos(\varphi)\\ 0\end{pmatrix} \tag{21.274c}$$

Nous avons vu que le produit mixte revient à mettre toutes les composantes dans une matrice. Ici nous avons donc

$$\frac{\partial\phi}{\partial r}\cdot\left(\frac{\partial\phi}{\partial\theta}\times\frac{\partial\phi}{\partial\varphi}\right) = \begin{vmatrix}\frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} & \frac{\partial z}{\partial r}\\ \frac{\partial x}{\partial\theta} & \frac{\partial y}{\partial\theta} & \frac{\partial z}{\partial\theta}\\ \frac{\partial x}{\partial\varphi} & \frac{\partial y}{\partial\varphi} & \frac{\partial z}{\partial\varphi}\end{vmatrix} \tag{21.275}$$

Cela est précisément le jacobien dont nous parlions plus haut.

21.12.5 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 une paramétrisation à une, deux ou trois variables.

(2) Dériver la paramétrisation 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.

21.12.6 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 $\sigma\colon [a,b]\to\mathbb{R}^3$ et le champ de vecteurs F, nous avons donc

$$\int_\sigma F=\int F\cdot d\sigma=\int_a^b F\big(\sigma(t)\big)\cdot\sigma'(t)dt. \tag{21.276}$$

21.12.7 Flux d'un champ de vecteurs

Si la surface $S\subset\mathbb{R}^3$ est paramétrée par

$$\begin{aligned}\phi\colon D\subset\mathbb{R}^2&\to\mathbb{R}^3\\(u,v)&\mapsto\phi(u,v),\end{aligned} \tag{21.277}$$

et si F est un champ de vecteurs, alors on a

$$\int_S F=\int_S F\cdot dS=\iint_D F\big(\phi(u,v)\big)\cdot\left(\frac{\partial\phi}{\partial u}\times\frac{\partial\phi}{\partial v}\right)dudv. \tag{21.278}$$

21.12.8 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

$$\frac{\partial\phi}{\partial u}\times\frac{\partial\phi}{\partial v} \tag{21.279}$$

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.

21.12.9 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

$$T_u\times T_v. \tag{21.280}$$

Mais le vecteur $T_v\times T_u$ est tout aussi normal ! Il n'y a pas a 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 [13], les deux réponses sont correctes.

Un exemple de ce problème est donné dans l'exemple 21.81.

Exemple 21.81

Calculer le flux du champ de vecteurs

$$F(x,y,z) = e_x \tag{21.281}$$

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 ».

Une paramétrisation du cylindre autour de l'axe z est

$$\phi(\theta,z) = \begin{pmatrix} R\cos\theta \\ R\sin\theta \\ z \end{pmatrix}. \tag{21.282}$$

Les vecteurs tangents sont

$$T_\theta = \begin{pmatrix} -R\sin\theta \\ R\cos\theta \\ 0 \end{pmatrix}, T_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{21.283}$$

Le vecteur normal est donc

$$T_\theta \times T_z = R\cos(\theta)e_x + R\sin(\theta)e_y. \tag{21.284}$$

C'est un vecteur dirigé vers l'extérieur.

Le champ de vecteurs considéré est constant : $F(\theta,z) = e_x$. Nous avons donc

$$F(\theta,z) \cdot (T_\theta \times T_z) = R\cos(\theta) \tag{21.285}$$

et le flux vaut

$$\Phi = \int_0^{2\pi} d\theta \int_0^h R\cos(\theta)dz = 0. \tag{21.286}$$

En ce qui concerne les couvercles haut au bas, ils sont paramétrés par

$$\phi_1(r,\theta) = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ h \end{pmatrix}, \phi_2(r,\theta) = \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \\ 0 \end{pmatrix}. \tag{21.287}$$

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. Une paramétrisation est

$$\phi(\theta,x) = \begin{pmatrix} x \\ R\cos(\theta) \\ R\sin(\theta) \end{pmatrix}, \tag{21.288}$$

et le vecteurs tangents sont

$$T_\theta = \begin{pmatrix} 0 \\ -R\sin\theta \\ R\cos\theta \end{pmatrix}, T_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{21.289}$$

Le vecteur normal est alors donné par

$$T_\theta \times T_x = R\cos(\theta)e_y + R\sin(\theta)e_z. \tag{21.290}$$

13. Il faudrait définir ce qu'est une surface *orientable* et choisir une orientation.

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 la paramétrisation

$$\phi_1(r, \theta) = \begin{pmatrix} 0 \\ r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \tag{21.291}$$

Le vecteur normal serait $T_r \times T_\theta = re_x$, et le flux

$$\Phi = \int_0^{2\pi} d\theta \int_0^R r\, dr = \pi R^2. \tag{21.292}$$

Le second couvercle est donné par

$$\phi_2(r, \theta) = \begin{pmatrix} h \\ r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}. \tag{21.293}$$

Le vecteur normal est encore re_x, et le flux est à nouveau πR^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 le 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 $-re_x$. Le premier flux vaudra donc

$$-\pi R^2, \tag{21.294}$$

de telle sorte que le flux total sera nul. △

21.13 Formes différentielles exactes et fermées

Nous avons déjà parlé de formes différentielles et de leurs intégrales sur un chemin dans la section 13.19.

Définition 21.82.
*La forme différentielle ω est **exacte** s'il existe une fonction f telle que $\omega = df$; elle est dite **fermée** si $d\omega = 0$.*

Dire que la forme différentielle $\omega = f dx + g dy$ est fermée, c'est dire que

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}. \tag{21.295}$$

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 21.83.
Si ω est une 1-forme exacte de classe C^1, alors ω est fermée.

Démonstration. Le fait que ω soit exacte implique l'existence d'une fonction f telle que $\omega = df$, c'est-à-dire

$$\omega_x = \sum_i a_i(x) dx_i = \sum_i \frac{\partial f}{\partial x_i}(x) dx_i, \tag{21.296}$$

c'est-à-dire que $a_i(x) = \frac{\partial f}{\partial x_i}(x)$. L'hypothèse que ω est C^1 implique que f est C^2, et donc que nous pouvons inverser l'ordre de dérivation pour les dérivées secondes $\partial^2_{ij} f = \partial^2_{ji} f$. Nous pouvons donc faire le calcul suivant :

$$\frac{\partial a_i}{\partial x_j} = \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i} = \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j} = \frac{\partial a_j}{\partial x_i}, \tag{21.297}$$

ce qu'il fallait démontrer. □

La réciproque est vraie sur un ouvert simplement connexe.

Théorème 21.84.
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 $\mathbb{R}^n$.

Théorème 21.85.
Soit $D \subset \mathbb{R}^2$, une ouvert étoilé, et ω, une 1-forme fermée de classe C^1. Alors ω est exacte.

Démonstration. Soit $D \subset \mathbb{R}^2$, un ouvert étoilé par rapport à l'origine. Soient $f\colon D \to \mathbb{R}$, $g\colon D \to \mathbb{R}$, des fonctions de classe C^1 telles que

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial x} \tag{21.298}$$

sur D, et

$$F(x,y) = \int_0^1 \big[f(tx,ty)x + g(tx,ty)y\big]dt \tag{21.299}$$

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 18.25, et nous pouvons dériver sous le signe intégral.

Nous calculons, en utilisant la règle de dérivation de fonctions composées

$$\begin{aligned}\frac{\partial F}{\partial x}(x,t) &= \int_0^1 \left[f\frac{\partial f}{\partial x}(tx,ty)x + f(tx,ty) + t\frac{\partial g}{\partial x}(tx,ty)y\right]dt \\ &= \int_0^1 \left[t\Big(x\frac{\partial f}{\partial x}(tx,ty) + y\frac{\partial f}{\partial y}(tx,ty)\Big) + f(tx,ty)\right]dt\end{aligned} \tag{21.300}$$

où nous avons utilisé l'hypothèse $\partial_y f = \partial_x g$. Ce qui se trouve dans la parenthèse n'est autre que $\partial_t\big(f(tx,ty)\big)$, plus précisément, si nous posons $\mathcal{F}(x,y,t) = f(tx,ty)$, nous avons

$$\frac{\partial \mathcal{F}}{\partial t}(x,y,t) = x\frac{\partial f}{\partial x}(tx,ty) + y\frac{\partial f}{\partial y}(tx,ty). \tag{21.301}$$

En recopiant le résultat (21.300) en termes de $\mathcal{F}$, nous avons

$$\begin{aligned}\frac{\partial F}{\partial x}(x,t) &= \int_0^1 \left(t\frac{\partial \mathcal{F}}{\partial t}(x,y,t) + \mathcal{F}(x,y,t)\right)dt \\ &= \int_0^1 \partial_t\big(t\mathcal{F}(x,y,t)\big)dt \\ &= \big[f\mathcal{F}(x,y,t)\big]_0^1 \\ &= \mathcal{F}(x,y,1) \\ &= f(x,y).\end{aligned} \tag{21.302}$$

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

$$\frac{\partial F}{\partial x} = f, \quad \text{et} \quad \frac{\partial F}{\partial y} = g. \tag{21.303}$$

En ayant prouvé cela, nous avons prouvé que si $\omega = fdx + gdy$ avec $\partial_y f = \partial_x g$, alors $\omega = dF$ où F est définie par (21.299). □

Démonstration alternative du théorème 21.85. 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

$$F(x, y) = \int_0^1 \big(x\mathcal{F}(x, y, t) + y\mathcal{G}(x, y, t)\big)dt, \tag{21.304}$$

et

$$\begin{aligned} \frac{\partial \mathcal{F}}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = t\frac{\partial f}{\partial u}, \\ \frac{\partial \mathcal{G}}{\partial x} &= t\frac{\partial g}{\partial u}. \end{aligned} \tag{21.305}$$

Ainsi,

$$\begin{aligned} \frac{\partial F}{\partial x} &= \int_0^1 \left(x\frac{\partial \mathcal{F}}{\partial x} + \mathcal{F} + y\frac{\partial G}{\partial x} \right) dt \\ &= \int_0^1 \left(xt\frac{\partial f}{\partial u} + \mathcal{F} + yt\frac{\partial g}{\partial u} \right) dt \\ &= \int_0^1 \left[t\left(x\frac{\partial f}{\partial u} + y\frac{\partial f}{\partial v} \right) + \mathcal{F} \right] dt. \end{aligned} \tag{21.306}$$

où nous avons utilisé le fait que, par hypothèse, $\frac{\partial g}{\partial u} = \frac{\partial f}{\partial v}$. Nous calculons par ailleurs que

$$\frac{\partial F}{\partial t} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial t} = x\frac{\partial f}{\partial u} + y\frac{\partial f}{\partial v}. \tag{21.307}$$

Donc, nous avons

$$\frac{\partial F}{\partial x} = \int_0^1 \left(t\frac{\partial \mathcal{F}}{\partial t} + \mathcal{F} \right) dt = \int_0^1 \frac{\partial}{\partial t}(t\mathcal{F})dt. \tag{21.308}$$

Par conséquent,

$$\frac{\partial F}{\partial x} = [t\mathcal{F}]_0^1 = \mathcal{F}(x, y, 1) = f(x, y). \tag{21.309}$$

Le même genre de calculs fournit $\frac{\partial F}{\partial y} = g(x, y)$. □

21.14 Théorème d'Abel angulaire

Théorème 21.86 (Abel angulaire[43])**.**
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

$$\Delta_{\theta_0} = \{z = 1 - \rho e^{i\varphi} \text{ tel que } \rho > 0, \varphi \in [\theta_0, \theta_0], |z| < 1\}. \tag{21.310}$$

Nous supposons de plus que $\sum_n a_n$ converge. Alors

$$\lim_{\substack{z \to 1 \\ z \in \Delta_0}} f(z) = \sum_{k=0}^{\infty} a_k. \tag{21.311}$$

Démonstration. Le résultat de ce théorème est que l'on peut calculer la limite $z \to 1$ avec des chemins contenus dans un domaine de la forme de celui dessiné à la figure 21.6.

De façon très classique nous posons

$$S = \sum_{k=0}^{\infty} a_k \quad S_n = \sum_{k=0}^{n} a_k, \tag{21.312}$$

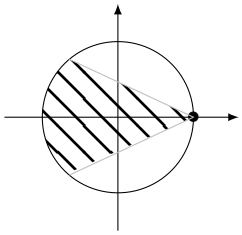


FIGURE 21.6 – La zone dans laquelle peut être le chemin qui va vers $z = 1$.

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 Δ_{θ_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 :

$$\sum_{n=0}^{N} a_n z^n - S_N = \sum_{n=1}^{N} a_n(z^n - 1) \tag{21.313a}$$

$$= \sum_{n=1}^{N} (R_{n-1} - R_n)(z^n - 1) \tag{21.313b}$$

$$= R_0(z-1) + \sum_{n=1}^{N-1} R_n(z^{n+1} - 1 - z^n + 1) + R_N(z^N - 1) \tag{21.313c}$$

$$= R_0(z-1) + \sum_{n=1}^{N-1} R_n z^n(z-1) + R_N(z^N - 1) \tag{21.313d}$$

$$= (z-1)\sum_{n=0}^{N-1} R_n z^n + R_N(z^N - 1). \tag{21.313e}$$

Cela est valable pour tout N et $|z| < 1$. Nous avons donc

$$\sum_{n=0}^{N} a_n z^n - S_N = (z-1)\sum_{n=0}^{N-1} R_n z^n + R_N(z^N - 1). \tag{21.314}$$

Par hypothèse nous avons $\lim_{N\to\infty} R_N = 0$. 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 $|z| < 1$:

$$f(z) - S = (z-1)\sum_{n=0}^{\infty} R_n z^n. \tag{21.315}$$

Nous voudrions étudier le comportement de la différence $f(z) - S$ lorsque z tend vers 1. Pour cela nous nous fixons $\epsilon > 0$ et $N \geqslant 1$ tel que $|R_n| < \epsilon$ dès que $n \geqslant N$. Alors pour tout $|z| < 1$ nous avons

$$|f(z) - S| \leqslant |z-1| \left(\sum_{n=0}^{N} |R_n| \underbrace{|z^n|}_{\leqslant 1} + \sum_{n=N+1}^{\infty} \underbrace{|R_n|}_{\leqslant \epsilon} |z^n| \right) \tag{21.316a}$$

$$\leqslant |z-1| \sum_{n=0}^{N} |R_n| + \epsilon \frac{|z-1|}{1-|z|} \tag{21.316b}$$

où nous avons utilisé la somme de la série géométrique (12.142) 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 :

$$\frac{|z-1|}{1-|z|} = \frac{|z-1|(1+|z|)}{1-|z|^2}. \tag{21.317}$$

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 :

$$\frac{|z-1|}{1-|z|} = \frac{\rho(1+|z|)}{2\rho\cos(\varphi) - \rho^2} \tag{21.318a}$$
$$= \frac{1+|z|}{2\cos(\varphi) - \rho} \tag{21.318b}$$
$$\leqslant \frac{2}{2\cos(\varphi) - \rho} \tag{21.318c}$$
$$\leqslant \frac{2}{2\cos(\varphi) - \cos(\theta_0)} \tag{21.318d}$$
$$\leqslant \frac{2}{2\cos(\theta_0) - \cos(\theta_0)} \tag{21.318e}$$
$$= \frac{2}{\cos(\theta_0)}. \tag{21.318f}$$

Quelques justifications.

— Vu que nous avons dans l'idée de faire $\rho \to 0$ nous supposons que $\rho < \cos(\theta_0)$.

— Nous avons $\cos(\varphi) > \cos(\theta_0)$ parce que z est dans Δ_{θ_0}.

Nous avons donc, pour tout $z \in \Delta_{\theta_0}$ que

$$|f(z) - S| \leqslant |z-1| \sum_{n=0}^{N} |R_n| + \epsilon \frac{2}{\cos(\theta_0)}. \tag{21.319}$$

Il suffit de prendre ρ assez petit pour que

$$|z-1| \sum_{n=0}^{N} |R_n| < \epsilon \tag{21.320}$$

et nous avons

$$|f(z) - S| \leqslant \epsilon \left(1 + \frac{2}{\cos(\theta_0)}\right). \tag{21.321}$$

Nous avons donc bien $\lim_{\substack{z\to 1 \\ z\in\Delta_0}} f(z) = S$, comme nous le voulions. □

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 21.87
Nous considérons la série entière $\sum_{n=0}^{\infty}(-1)^n z^n$ qui converge[14] vers

$$f(z) = \frac{1}{1+z} \tag{21.322}$$

sur $B(0,1)$. De plus nous avons

$$\lim_{\substack{z\to 1 \\ |z|<1}} \frac{1}{1+z} = \frac{1}{2}. \tag{21.323}$$

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 es coefficients $\sum_n (-1)^n$ ne converge pas. △

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.

14. C'est la série géométrique de raison $-z$.

Théorème 21.88 (Théorème taubérien faible[43]).
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 $\lim_{\substack{x\to 1\\ x\in]-1,1[}} f(x) = S$.

(2) $\lim_{n\to\infty} na_n = 0$.

Alors la série $\sum_{n=0}^{\infty} a_n$ converge et vaut S.

Démonstration. Nous notons $S_n = \sum_{k=0} a_k$ et $M = \sup_{k\geqslant 1} k|a_k|$, qui est fini par hypothèse. Pour $x \in]0,1[$ et $n \geqslant 0$ nous avons

$$S_n - f(x) = \sum_{k=1}^{n} a_k - \sum_{k=1}^{n} a_k x^k - \sum_{k=n+1}^{\infty} a_k x^k = \sum_{k=1}^{n} a_k(1-x^k) - \sum_{k=n+1}^{\infty} a_k x^k. \tag{21.324}$$

Nous utilisons la série géométrique sous la forme $1 - x^k = (1-x)\sum_{i=0}^{n} x^i$ pour écrire

$$S_n - f(x) = \sum_{k=1}^{n} a_k(1-x) \underbrace{\sum_{i=0}^{k-1} x^i}_{\leqslant k} - \sum_{k=n+1}^{\infty} a_k x^k \tag{21.325a}$$

$$\leqslant \sum_{k=1}^{n} k a_k(1-x) - \sum_{k=n+1}^{\infty} a_k x^k, \tag{21.325b}$$

donc en passant à la norme

$$\big|S_n - f(x)\big| \leqslant (1-x)Mn + \sum_{k=n+1} |a_k| x^k \tag{21.326a}$$

$$\leqslant (1-x)Mn + \sum_{k=n+1}^{\infty} \underbrace{\frac{k}{n}|a_k|}_{\leqslant M/n} x^k \tag{21.326b}$$

$$\leqslant (1-x)Mn + \frac{M}{n}\sum_{k=n+1}^{\infty} x^k \tag{21.326c}$$

$$\leqslant (1-x)Mn + \frac{M}{n}\frac{1}{1-x}. \tag{21.326d}$$

Ce que nous cherchons à étudier est le comportement $x \to 1$ et montrer que $S_n \to S$, ce qui nous incite à calculer $|S_n - f(1-\frac{\epsilon}{n})|$ avec $0 < \epsilon < 1$:

$$\big|S_n - f\big(1-\frac{\epsilon}{n}\big)\big| \leqslant \epsilon M + \epsilon. \tag{21.327}$$

Nous choisissons N_1 tel que $\frac{M}{n} \leqslant \epsilon^2$ dès que $n \geqslant N_1$. En sus nous savons que

$$\lim_{\epsilon\to 0} f(1-\epsilon) = S. \tag{21.328}$$

Nous choisissons N_2 de telle sorte à avoir

$$\left|f\left(1-\frac{\epsilon}{n}\right) - S\right| < \epsilon, \tag{21.329}$$

et en prenant $n \geqslant \max(N_1, N_2)$ nous avons

$$|S_n - S| \leqslant \left|S_n - f\left(1-\frac{\epsilon}{n}\right)\right| + \left|f\left(1-\frac{\epsilon}{n}\right) - S\right| \leqslant \epsilon M + 2\epsilon. \tag{21.330}$$

Il suffit de choisir ϵ suffisamment petit (en particulier pour que ϵM soit petit) pour montrer que $|S_n - S|$ est borné par un nombre arbitrairement petit. □

21.14.1 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 21.89.
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 $f\colon E \to \overline{\mathbb{R}}$. Supposons qu'il existe une fonction g intégrable sur E telle que pour tout k,

$$|f(x)| \leqslant g(x) \tag{21.331}$$

pour tout $x \in E$. Alors f est intégrable sur E et

$$\int_E f = \lim_{k\to\infty} \int_E f_k. \tag{21.332}$$

21.14.2 Intégrale en dimension un

Proposition 21.90 (Critère de comparaison)**.**
Soit f mesurable sur $]a,\infty[$ et bornée sur tout $]a,b]$, et supposons qu'il existe un $X_0 \geqslant a$, tel que sur $]X_0,\infty[$,

$$|f(x)| \leqslant g(x) \tag{21.333}$$

où $g(x)$ est intégrable. Alors $f(x)$ est intégrable sur $]a,\infty[$.

Corollaire 21.91 (Critère d'équivalence)**.**
Soient f et g des fonctions mesurables et positives ou nulles sur $]a,\infty[$, bornées sur tout $]a,b]$, telles que

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = L \tag{21.334}$$

existe dans $\bar{\mathbb{R}}$.

(1) Si $L \neq \infty$ et $\int_a^\infty g(x)$ 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,

Corollaire 21.92 (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

$$L(\alpha) = \lim_{x\to\infty} x^\alpha f(x), \tag{21.335}$$

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 $\alpha \leqslant 1$ et $L(\alpha) \neq 0$, alors $\int_a^\infty f(x)dx$ n'existe pas.

Corollaire 21.93.
Soit $f\colon]a,b] \to \mathbb{R}$ une fonction mesurable, positive ou nulle, et bornée sur $[a+\epsilon,b]$ $\forall \epsilon > 0$. Si $\lim_{x\to a}(x-a)^\alpha f(x) = L$ existe, alors

(1) Si $\alpha < 1$ et $L \neq \infty$, alors $\int_a^b f(x)dx$ existe,

(2) Si $\alpha \geqslant 1$ et $L \neq 0$, alors $\int_a^b f(x)dx$ n'existe pas.

21.14.3 Intégrales convergentes

Définition 21.94.
Soit f, une fonction mesurable sur $[a,\infty[$, bornée sur tout intervalle $[a,b]$. On dit que l'intégrale

$$\int_a^\infty f(x)dx \tag{21.336}$$

converge *si la limite*

$$\lim_{X\to\infty}\int_a^X f \tag{21.337}$$

existe et est finie.

21.14.4 La méthode de Rothstein-Trager

Mes sources pour parler d'intégration de fractions rationnelles : [355].

Théorème 21.95 (Rothstein-Trager[356])**.**
Soient $P,Q\in\mathbb{Q}[X]$ premiers entre eux avec $\operatorname{pgcd}(P,Q)=1$ et $\deg(P)<\deg(Q)$. Nous supposons que Q est unitaire et sans facteurs carrés. Supposons que nous puissions écrire, dans un extension $\mathbb{K}$ de $\mathbb{Q}$ la primitive de P/Q de la façon suivante :

$$\int\frac{P}{Q}=\sum_{i=1}^{n}c_i\ln(P_i) \tag{21.338}$$

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

$$R(Y)=\operatorname{res}_X(P-YQ',Q)\in\mathbb{K}[Y] \tag{21.339}$$

et

$$P_i=\operatorname{pgcd}(P-c_iQ',Q). \tag{21.340}$$

Démonstration. Nous posons

$$U_i=\prod_{j\neq i}P_j. \tag{21.341}$$

Question de division Ensuite nous dérivons formellement l'équation (21.338) et nous multiplions les deux côtés du résultat par $\prod_{j=1}^{n}P_j$:

$$P\prod_{j=1}^{n}P_j=Q\sum_{i=1}^{n}c_i\frac{P'i}{P_i}\prod_{j=1}^{n}P_j=Q\sum_{i=1}^{n}c_iP_i'U_i. \tag{21.342}$$

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,

$$Q\mid\prod_{j=1}^{n}P_j \tag{21.343}$$

par le théorème de Gauss 6.43.

Une seconde chose que nous tirons de (21.342) est que P_j divise $Q\sum_{i=1}^{n}c_iP_i'U_i$. 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

$$P_j\mid Qc_jP_J'U_j \tag{21.344}$$

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

$$P_j\mid Q \tag{21.345}$$

pour tout $1\leqslant j\leqslant n$. 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

$$\prod_{j=1}^{n}P_j\mid Q. \tag{21.346}$$

Or nous avions déjà prouvé la division contraire. Du fait que les deux polynômes sont unitaires nous en déduisons qu'ils sont en réalité égaux :

$$Q = \prod_{j=1}^{n} P_j. \tag{21.347}$$

Nous pouvons simplifier les deux membres de (21.342) par cela :

$$P = \sum_{i=1}^{n} c_i P_i' U_i. \tag{21.348}$$

Encore un peu de division En dérivant (21.347) nous trouvons

$$Q' = \sum_{j=1}^{n} P_j' U_j, \tag{21.349}$$

et en écrivant P sous sa forme (21.348),

$$P - c_i Q' = \sum_{j=1}^{n} c_j P_j' U_j - \sum_{j=1}^{n} c_i P_j' U_j = \sum_{j=1}^{n} (c_j - c_i) P_j' U_j. \tag{21.350}$$

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

$$P_i \mid P - c_i Q'. \tag{21.351}$$

Un pgcd pour continuer Nous montrons à présent que $P_i = \operatorname{pgcd}(P - c_i Q', Q)$. Pour cela nous utilisons la multiplicativité du PGCD lorsque les facteurs sont premiers entre eux :

$$\operatorname{pgcd}(P - c_i Q', Q) = \operatorname{pgcd}(P - c_i Q', \prod_{j=1}^{n} P_j) = \prod_{j=1}^{n} \operatorname{pgcd}(P - c_i Q', P_j). \tag{21.352}$$

Nous remplaçons $P - c_i Q'$ par son expression (21.350) et nous écrivons un des facteurs du produit :

$$\operatorname{pgcd}(P - c_i Q', P_j) = \operatorname{pgcd}(\sum_{k=1}^{n} (c_k - c_i) P_k' U_k, P_j) \tag{21.353}$$

Le polynôme P_j divise tous les U_k sauf celui avec $k = j$. Donc le lemme 6.49(1) nous permet de dire

$$\operatorname{pgcd}(P - c_i Q', P_j) = \operatorname{pgcd}\big((c_j - c_i) P_j' U_j, P_j\big) = \begin{cases} 1 & \text{si } i \neq j \\ P_j & \text{si } i = j. \end{cases} \tag{21.354}$$

La seconde ligne provient du fait que nous ayons déjà montré que $P_j \mid P - c_j Q'$. En fin de compte,

$$\operatorname{pgcd}(P - c_i Q', Q) = P_i. \tag{21.355}$$

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 11.60 nous dit que le résultant

$$\operatorname{res}_X(P - c_i Q', Q) = 0. \tag{21.356}$$

Donc les c_i sont des racines du polynôme (en Y)

$$R(Y) = \operatorname{res}_X(P - YQ', Q). \tag{21.357}$$

Nous n'avons pas prouvé qu'ils étaient *toutes* les racines [15].

15. 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.

Toutes les racines Nous allons maintenant montrer que les c_i étaient toutes les racines imaginables du polynôme (21.357) 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 (21.338). Étant donné que c est racine du résultat, les polynômes $P-cQ'$ et Q ont un PGCD non trivial, c'est-à-dire non constant. Donc

$$\text{pgcd}(P - cQ', Q) = s \in \hat{\mathbb{K}}[X] \tag{21.358}$$

est un polynôme non constant. Si T un facteur irréductible de S, alors T divise $P - cQ'$ 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 [16]. Soit P_{i_0} celui qui est divisé par T. La relation (21.350) dans ce contexte donne :

$$P - cQ' = \sum_{j=1}^{n} (c_j - c)P'_j U_j \tag{21.359}$$

Le polynôme T divise tous les U_j avec $j \neq i_0$, mais comme en plus il divise $P - cQ'$, il divise aussi le dernier terme de la somme :

$$T \mid (c_{i_0} - c)P'_{i_0} U_{i_0}. \tag{21.360}$$

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.

□

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 $P_p = \text{pgcd}(P - c_i Q', Q)$.

(3) Écrire la réponse :

$$\int \frac{P}{Q} = \sum_{i=1}^{n} c_i \ln(P_i). \tag{21.361}$$

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 a priori pas pire à factoriser que Q lui-même [17]. 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 [18].

Exemple 21.96([356])
Prenons la fraction rationnelle $\frac{x}{x^2-3}$. L'intégration via les fractions simples est :

$$\int \frac{x}{x^2-3}dx = \frac{1}{2}\int \frac{1}{x-\sqrt{3}} + \frac{1}{2}\int \frac{1}{x+\sqrt{3}} = \frac{1}{2}\ln(x-\sqrt{3}) + \frac{1}{2}\ln(x+\sqrt{3}) = \frac{1}{2}\ln(x^2-3). \tag{21.362}$$

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.

16. On ne peut pas diviser deux trucs qui sont premiers entre eux ; c'est une question de cohérence, madame !
17. C'est de la factorisation de Q qu'on a besoin pour utiliser la méthode de décomposition en fractions simples.
18. J'imagine que pour un ordinateur, c'est plus facile d'éviter les extensions.

D'abord

$$\begin{aligned} R(Y) &= \operatorname{res}_X(X - 2YX, X^2 - 3) && (21.363a)\\ &= \operatorname{res}_X\big((1-2Y)X, X^2 - 3\big) && (21.363b)\\ &= \det\begin{pmatrix} 1-2Y & 0 & 0\\ 0 & 1-2Y & 0\\ 1 & 0 & -3 \end{pmatrix} && (21.363c)\\ &= (1-2Y)\big(-3(1-2Y)\big) && (21.363d)\\ &= -3(1-2Y)^2, && (21.363e) \end{aligned}$$

dont les solutions sont faciles : il n'y a que la racine double $y = \frac{1}{2}$. La somme (21.338) sera donc réduite à un seul terme avec $c_1 = \frac{1}{2}$. Nous calculons P_1 :

$$P_1 = \operatorname{pgcd}(X - \frac{1}{2}2X, X^2 - 3) = \operatorname{pgcd}(0, X^2 - 3) = X^2 - 3, \tag{21.364}$$

et par conséquent

$$\int \frac{X}{X^2 - 3} = \frac{1}{2}\ln(X^2 - 3). \tag{21.365}$$

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 $\mathbb{Q}[X]$.

Remarque 21.97.
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

$$Q = Q_1 Q_2^2 Q_3^3 \dots Q_r^r, \tag{21.366}$$

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.

Exemple 21.98([357])
À calculer : $\int \frac{1}{x^3+x}$. La décomposition en fractions simples donne :

$$\frac{1}{x^3 + x} = \frac{1}{x} - \frac{1/2}{x - i} - \frac{1/2}{x + i}. \tag{21.367}$$

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 27.67 au prix d'un peu de sang. Nous verrons dans l'exemple 27.68 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) :

$$\begin{aligned} P(y) &= \operatorname{res}_X\big(-3yX^2 - y + 1, X^3 - X\big) && (21.368a)\\ &= \det\begin{pmatrix} -3y & 0 & 1-y & 0 & 0\\ 0 & -3y & 0 & 1-y & 0\\ 0 & 0 & -3y & 0 & 1-y\\ 1 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} && (21.368b)\\ &= -(y-1)^2(2y+1)^2 && (21.368c) \end{aligned}$$

```
----------------------------------------------------------------------
| 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 :

$$P_1 = \text{pgcd}(-3X^2, X^3 + X) = X \tag{21.369}$$

et

$$P_2 = \text{pgcd}(\frac{3}{2}X^2 + \frac{3}{2}, X^3 + X) = X^2 + 1, \tag{21.370}$$

et finalement

$$\int \frac{1}{X^3 + X} = \ln(X) - \frac{1}{2}\ln(X^2 + 1). \tag{21.371}$$

△

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.

Exemple 21.99([357])
Nous voulons calculer

$$\int \frac{1}{X^2 + 1}. \tag{21.372}$$

Nous posons donc $P = 1$ et $Q = X^2 + 1$. Le résultant à calculer est

$$P(y) = \text{res}_X(-2yX + 1, X^2 + 1) = \det \begin{pmatrix} -2y & 1 & 0 \\ 0 & -2y & 1 \\ 1 & 0 & 1 \end{pmatrix} = 4y^2 + 1. \tag{21.373a}$$

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 $X^2 + 1 = (X + i)(X - i) = i(-iX + 1)(X - 1)$ nous avons

$$P_1 = \text{pgcd}(-iX + 1, X^2 + 1) = X + i. \tag{21.374}$$

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,

$$P_2 = \text{pgcd}(1 + iX, X^2 + 1) = X - i. \tag{21.375}$$

Au final nous écrivons

$$\int \frac{1}{X^2 + 1} = \frac{i}{2}\ln(X + i) - \frac{i}{2}\ln(X - i). \tag{21.376}$$

△

Remarque 21.100.
Tout cela est si nous voulons absolument écrire la primitive avec des logarithmes de polynômes. Pour celui de l'exemple 21.99, nous avons trouvé

$$\int \frac{1}{X^2 + 1} = \frac{i}{2}\ln(X + i) - \frac{i}{2}\ln(X - i). \tag{21.377}$$

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 [19] compatibles. En effet, afin de tomber juste, nous allons prendre la primitive

$$f(x) = \frac{i}{2}\ln(ix-1) - \frac{i}{2}\ln(ix+1) \tag{21.378}$$

au lieu de (21.377). 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,

$$f(x) = -\frac{1}{2}\Big(\arctan(-x) - \arctan(x)\Big) = \arctan(x). \tag{21.379}$$

21.15 Rappel sur les intégrales usuelles

Soit une fonction

$$\begin{aligned} f\colon [a,b] \subset \mathbb{R} &\to \mathbb{R}^+ \\ x &\mapsto f(x). \end{aligned} \tag{21.380}$$

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 21.7.

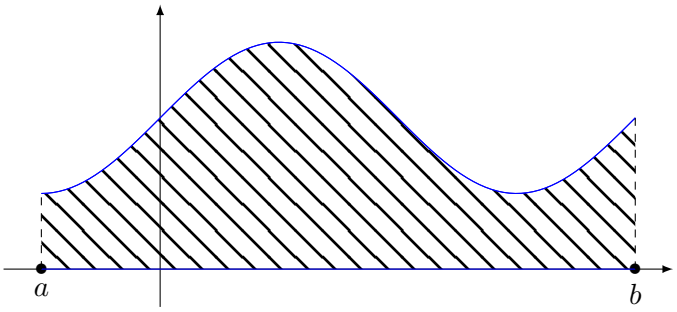


FIGURE 21.7 – L'intégrale de f entre a et b représente la surface sous la fonction.

Définition 21.101.
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 21.102 (Théorème fondamental du caclul intégral)**.**
Si f est une fonction positive et continue, et si F est une primitive de f, alors

$$\int_a^b f(x)dx = F(b) - F(a). \tag{21.381}$$

Remarque 21.103.
Si f est une fonction continue par morceaux, l'intégrale de f se calcule comme la somme des

19. Si on croit que la mathématique est cohérente.

intégrales de ses morceaux. Plus précisément si nous avons $a = x_0 < x_1 < \ldots < x_n = b$ et si f est continue sur $]x_i, x_{i+1}[$ pour tout i, alors nous posons

$$\int_a^b f(x)dx = \int_{x_0}^{x_1} f(x)dx + \int_{x_1}^{x_2} f(x)dx + \cdots + \int_{x_{n-1}}^{n_n} f(x)dx. \tag{21.382}$$

Sur chacun des morceaux, l'intégrale se calcule normalement en passant par une primitive.

21.16 Intégrales le long de chemins

21.16.1 Circulation d'un champ de vecteur

Définition 21.104.
Soit $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs et un chemin $\sigma\colon [a,b] \to \mathbb{R}^3$. On appelle ***circulation*** *de F le long du chemin σ le scalaire*

$$\int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt. \tag{21.383}$$

Il existe de nombreuses notations pour cela ; entre autres :

$$\int_\sigma F = \int_\sigma F \cdot ds. \tag{21.384}$$

En physique, la circulation de la force le long d'un chemin est la travail de la force.

Exemple 21.105
À la surface de la Terre, le champ de gravitation est donné par

$$G(x,y,z) = -mg\begin{pmatrix}0\\0\\1\end{pmatrix}. \tag{21.385}$$

Si nous considérons un mobile qui monte à vitesse constante jusqu'à la hauteur h, c'est-à-dire le chemin

$$\sigma(t) = \begin{pmatrix}0\\0\\t\end{pmatrix} \tag{21.386}$$

avec $t \in [0,h]$. Le travail de la gravitation est alors donné par

$$W = \int_0^h G\big(\sigma(t)\big) \cdot \begin{pmatrix}0\\0\\1\end{pmatrix} = -mg\int_0^h \begin{pmatrix}0\\0\\1\end{pmatrix} \cdot \begin{pmatrix}0\\0\\1\end{pmatrix} = -mgh. \tag{21.387}$$

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. △

Exemple 21.106
Soit le chemin

$$\begin{aligned}\sigma\colon [0,2\pi] &\to \mathbb{R}^3\\ t &\mapsto \begin{pmatrix}\sin(t)\\ \cos(t)\\ t\end{pmatrix}.\end{aligned} \tag{21.388}$$

et le champ de vecteurs

$$F\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}x\\y\\z\end{pmatrix}. \tag{21.389}$$

La circulation de ce champ de vecteur le long de l'hélice σ est

$$\begin{aligned}\int_\sigma F\cdot ds &= \int_0^{2\pi}(F\circ\sigma)(t)\cdot\sigma'(t)dt\\ &= \int_0^{2\pi}\begin{pmatrix}\sin(t)\\ \cos(t)\\ t\end{pmatrix}\cdot\begin{pmatrix}\cos(t)\\ \sin(t)\\ 1\end{pmatrix}dt\\ &= \int_0^{2\pi}tdt\\ &= \left[\frac{t^2}{2}\right]_0^{2\pi}\\ &= 2\pi^2.\end{aligned}\tag{21.390}$$

△

Proposition 21.107.
La circulation d'un champ de vecteurs le long d'un chemin ne dépend pas de la paramétrisation. En d'autres termes, si σ_1 et σ_2 sont deux chemins équivalents, alors

$$\int_{\sigma_1}F=\int_{\sigma_2}F.\tag{21.391}$$

Démonstration. Soient deux chemins $\sigma_1\colon[a,b]\to\mathbb{R}^3$ et $\sigma_2\colon[c,d]\to\mathbb{R}^3$ équivalents, c'est-à-dire tels que

$$\sigma_1(t)=\sigma_2\big(\varphi(t)\big)\tag{21.392}$$

où $\varphi\colon[a,b]\to[c,d]$ strictement croissante. En utilisant le fait que $\sigma_1(t)=\varphi'(t)\sigma_2'\big(\varphi(t)\big)$, nous avons

$$\begin{aligned}\int_{\sigma_1}F\cdot ds &= \int_a^b F\big(\sigma_1(t)\big)\cdot\sigma_1'(t)dt\\ &= \int_a^b F\Big(\sigma_2\big(\varphi(t)\big)\Big)\cdot\sigma_2'\big(\varphi(t)\big)\varphi'(t)dt\\ &= \int_{\varphi(a)}^{\varphi(b)}F\big(\sigma_2(s)\big)\cdot\sigma_2(s)ds\\ &= \int_c^d F\big(\sigma_2(s)\big)\cdot\sigma_2'(s)ds\\ &= \int_{\sigma_2}F\cdot ds.\end{aligned}\tag{21.393}$$

où nous avons effectué le changement de variables $s=\varphi(t)$, $ds=\varphi'(t)dt$. □

Remarque 21.108.
Si σ_2 est le chemin opposé de σ, alors

$$\int_{\sigma_2}F=-\int_{\sigma_1}F.\tag{21.394}$$

21.17 Circulation d'un champ conservatif

Si nous avons une fonction scalaire $V\colon\mathbb{R}^3\to\mathbb{R}$, nous pouvons construire un champ de vecteur en prenant le gradient :

$$F(x)=\nabla V(x).\tag{21.395}$$

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 21.109.
Un champ de vecteurs $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ est un champ ***conservatif*** *s'il existe une fonction $V\colon \mathbb{R}^3 \to \mathbb{R}$ telle que*

$$F(x) = \nabla V(x). \tag{21.396}$$

Nous disons aussi parfois que le champ V 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 21.110.
Considérons une fonction $V\colon \mathbb{R}^3 \to \mathbb{R}$ (que nous appellerons potentiel*) et le champ de vecteur qui en dérive :*

$$F = \nabla V. \tag{21.397}$$

Alors

$$\int_\sigma F \cdot ds = V\big(\sigma(b)\big) - V\big(\sigma(a)\big). \tag{21.398}$$

Autrement dit, le travail nécessaires 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,

$$\int_\sigma F \cdot ds = \int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt. \tag{21.399}$$

Nous pouvons transformer l'intégrante de la façon suivante :

$$\begin{aligned} F\big(\sigma(t)\big) \cdot \sigma'(t) &= \nabla V\big(\sigma(t)\big) \cdot \sigma'(t) \\ &= \frac{\partial V}{\partial x}\big(\sigma(t)\big)\sigma_x'(t) + \frac{\partial V}{\partial y}\big(\sigma(t)\big)\sigma_y'(t) + \frac{\partial V}{\partial z}\big(\sigma(t)\big)\sigma_z'(t) \\ &= \frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big] \end{aligned} \tag{21.400}$$

où nous avons posé

$$\sigma(t) = \begin{pmatrix} \sigma_x(t) \\ \sigma_y(t) \\ \sigma_z(t) \end{pmatrix} \tag{21.401}$$

et utilisé à l'envers la formule de dérivation de fonction composée pour

$$\frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big] = \Big((V \circ \sigma)(t)\Big)'. \tag{21.402}$$

En remettant ces expressions dans l'intégrale (21.399),

$$\int_\sigma F \cdot ds = \int_a^b \frac{d}{dt}\Big[V\big(\sigma(t)\big)\Big]dt = V\big(\sigma(b)\big) - V\big(\sigma(a)\big). \tag{21.403}$$

□

Exemple 21.111
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 suivit : en ligne droite vertical, en diagonal, en hélice, ... △

Exemple 21.112
À plus grande échelle, le champ de gravitation est encore un champ qui dérive d'un potentiel. En coordonnées sphériques,

$$V(\rho,\theta,\varphi)=k\frac{m}{\rho} \tag{21.404}$$

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. △

Exemple 21.113
Soit le champ de vecteurs

$$F\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}y\\x\end{pmatrix} \tag{21.405}$$

et le chemin

$$\sigma(t)=\begin{pmatrix}t^4/4\\ \sin^3(t\frac{\pi}{2}).\end{pmatrix} \tag{21.406}$$

Nous voulons calculer la circulation de F le long du chemin σ 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

$$\int_\sigma F\cdot ds=V\big(\sigma(1)\big)-V\big(\sigma(0)\big)=V\big(\frac{1}{4},1\big)-V(0,0)=\frac{1}{4}-0=\frac{1}{4}. \tag{21.407}$$

Nous n'avons pas réellement calculé l'intégrale. △

21.18 Intégration de fonction à deux variables

21.18.1 Intégration sur un domaine rectangulaire

Soit une fonction positive

$$\begin{aligned}f\colon [a,b]\times[c,d]&\to\mathbb{R}^+\\(x,y)&\mapsto f(x,y).\end{aligned} \tag{21.408}$$

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

$$\{(x,y,z)\text{ tel que }(x,y)\in[a,b]\times[c,d],z\leqslant f(x,y)\}. \tag{21.409}$$

Théorème 21.114 (Théorème de Fubini)**.**
Soit une fonction $f\colon\mathbb{R}^2\to\mathbb{R}$ une fonction continue par morceaux sur $\mathcal{R}=[a,b]\times[c,d]$. Alors

$$\int_{\mathcal{R}}f(x,y)dxdy=\int_a^b\left[\int_c^d f(x,y)dy\right]dx=\int_c^d\left[\int_a^b f(x,y)dx\right]dy. \tag{21.410}$$

En pratique, nous utilisons le théorème de Fubini pour calculer les intégrales sur des rectangles.

Exemple 21.115
Nous voudrions intégrer la fonction $f(x,y)-4+x^2+y^2$ sur le rectangle de la figure 21.8.

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 :

$$\int_{E=[0,1]\times[0,2]}f=\int_{[0,1]}\left(\int_{[0,2]}(4-x^2-y^2)dy\right)dx. \tag{21.411}$$

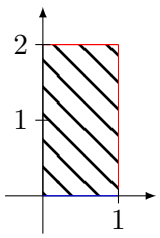


FIGURE 21.8 – Intégration sur un rectangle

Ces intégrales sont maintenant des intégrales usuelles qui s'effectuent en calculant des primitives :

$$\begin{aligned}\int_0^1\int_0^2(4-x^2-y^2)dy\,dx &= \int_0^1\left[4y-x^2y-\frac{y^3}{3}\right]_0^2dx\\ &= \int_0^1(8-2x^2-\frac{8}{3})dx\\ &= \left[\frac{16x}{3}-\frac{2x^3}{3}\right]_0^1\\ &= \frac{14}{3}.\end{aligned} \tag{21.412}$$

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
```

△

21.18.2 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 21.9.

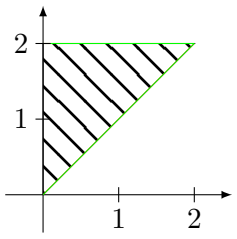
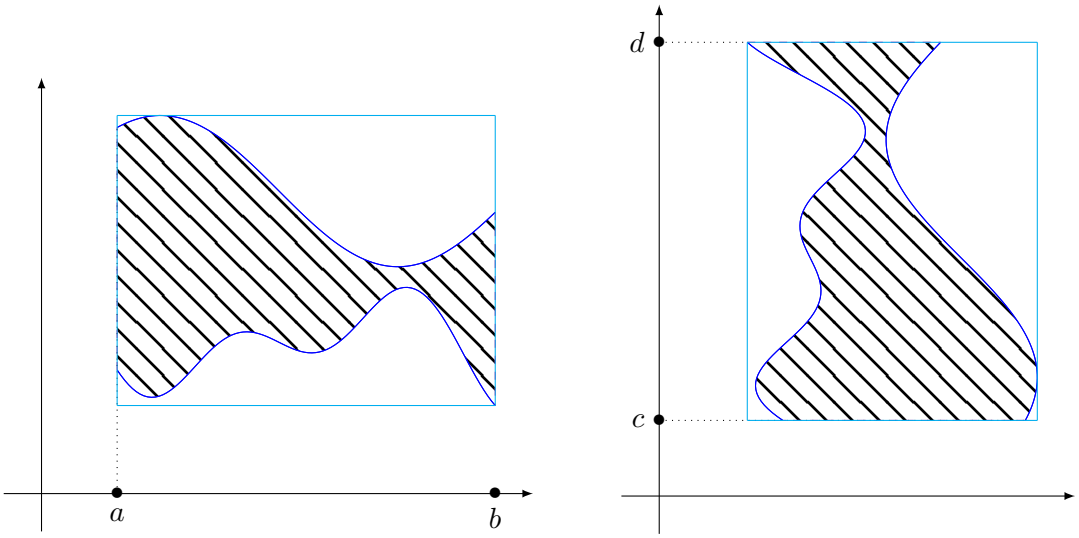


FIGURE 21.9 – Intégration sur un triangle

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 :

$$\int_{\text{triangle}}(x^2+y^2)dxdy=\int_0^2\left(\int_0^y(x^2+y^2)dx\right)dy. \tag{21.413}$$

Il existe principalement deux types de domaines non rectangulaires : les « horizontaux » et les « verticaux », voir figure 21.10.



(a) Un domaine horizontal.

(b) Un domaine vertical.

FIGURE 21.10 – 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

$$D = \{(x, y) \text{ tel que } x \in [a, b], \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\} \tag{21.414}$$

où φ_1 et φ_2 sont les deux fonctions qui bornent le domaine. Le domaine D est la région comprise entre les graphes de φ_1 et φ_2. Pour un tel domaine nous avons

$$\int_D f(x, y)dxdy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y)dy. \tag{21.415}$$

Les surfaces verticales sont de la forme

$$D = \{(x, y) \text{ tel que } y \in [c, d], \psi_1(y) \leqslant x \leqslant \psi_2(y)\} \tag{21.416}$$

où φ_1 et φ_2 sont les deux fonctions qui bornent le domaine. Le domaine D est la région comprise entre les graphes de φ_1 et φ_2. Dans ces cas nous avons

$$\int_D f = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} f(x, y)dx. \tag{21.417}$$

Proposition 21.116.
L'aire du domaine D vaut l'intégrable de la fonction $f(x, y) = 1$ sur D :

$$Aire(D) = \int_D dxdy. \tag{21.418}$$

Démonstration. Supposons que le domaine soit du type « horizontal ». En utilisant le théorème de Fubini avec $f(x, y) = 1$ nous avons

$$\int_D dxdy = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} dy \right] dx = \int_a^b \big[\varphi_2(x) - \varphi_1(x)\big]. \tag{21.419}$$

Cela représente l'aire sous φ_2 moins l'aire sous φ_1, et par conséquent l'aire contenue entre les deux. □

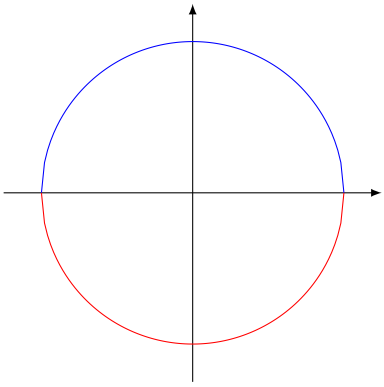


FIGURE 21.11 – En bleu, la fonction $\sqrt{r^2-x^2}$ et en rouge, la fonction $-\sqrt{r^2-x^2}$.

Exemple 21.117
Cherchons la surface du disque de centre $(0,0)$ et de rayon 1 dessinée à la figure 21.11.

Le domaine est donné par $\varphi_1(x) \leqslant y \leqslant \varphi_2(x)$ 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

$$A = \int_{-r}^{r} \big[\varphi_2(x) - \varphi_1(x)\big]dx = 2\int_{-r}^{r} \sqrt{r^2-x^2}dx = 4\int_{0}^{r} \sqrt{r^2-x^2}. \tag{21.420}$$

Nous effectuons le premier changement de variables $x = ru$, donc $dx = rdu$. En ce qui concerne les bornes, si $x = 0$, alors $u = 0$ et si $x = r$, alors $u = 1$. L'intégrale à calculer devient

$$A = 4\int_0^1 \sqrt{r^2 - r^2u^2}rdu = 4r^2\int_0^1 \sqrt{1-u^2}du. \tag{21.421}$$

Cette dernière intégrale se calcule en posant

$$\begin{aligned} &u = \sin(t) \quad && du = \cos(t)dt \\ &u = 0 && t = 0 \\ &u = 1 && t = \pi/2. \end{aligned} \tag{21.422}$$

Nous avons

$$A = 4r^2\int_0^{\pi/2} \sqrt{1-\sin^2(t)}\cos(t)dt = 4r^2\int_0^{\pi/2} \cos^2(t)dt. \tag{21.423}$$

En utilisant la formule $2\cos^2(x) = 1 + \cos(2x)$, nous avons

$$A = 4r^2\int_0^{\pi/2} \frac{1+\cos(2t)}{2}dt = \pi r^2. \tag{21.424}$$

△

21.18.3 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 19.14.

21.19 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 [20] à condition d'adapter les domaines d'intégration.

Exemple 21.118
Soit le domaine parallélépipédique rectangle

$$R = [0,1] \times [1,2] \times [0,4]. \tag{21.425}$$

Pour intégrer la fonction $f(x,y,z) = x^2 y \sin(z)$ sur R, nous faisons

$$\begin{aligned} I &= \int_R x^2 y \sin(z)\, dxdydz \\ &= \int_0^1 dx \int_1^2 dy \int_0^4 x^2 y \sin(z) dz \\ &= \int_0^1 dx \int_1^2 x^2 y(1-\cos(4)) dy \\ &= \int_0^1 \frac{3}{2}(1-\cos(4)) x^2 dx \\ &= \frac{1}{2}\big(1-\cos(4)\big). \end{aligned} \tag{21.426}$$

```
------------------------------------------------------------------------
| 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
```

△

Exemple 21.119
Soit D la région délimitée par le plan $x = 0$, $y = 0$, $z = 2$ et la surface d'équation

$$z = x^2 + y^2. \tag{21.427}$$

Cherchons à calculer $\int_D x\, dx\, dy\, dz$. Ici, un dessin indique que le volume considéré est $z \geqslant x^2 + y^2$. Il y a plusieurs façons de décrire cet ensemble. Une est celle-ci :

$$\begin{aligned} &z\colon 0 \to 2 \\ &x\colon 0 \to \sqrt{z} \\ &y\colon 0 \to \sqrt{z - x^2}. \end{aligned} \tag{21.428}$$

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

$$\int_0^2 dz \int_0^{\sqrt{z}} dx \int_0^{\sqrt{z-x^2}} f(x,y,z) dy. \tag{21.429}$$

```
------------------------------------------------------------------------
| Sage Version 4.6.1, Release Date: 2011-01-11                         |
```

20. En toute rigueur, cela n'est pas vrai, mais nous ne considérons seulement des cas où cela est autorisé.

```
| 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 $z - x^2 > 0$.

Une autre paramétrisation serait

$$\begin{aligned} x&\colon 0 \to \sqrt{2} \\ y&\colon 0 \to \sqrt{2-x^2} \\ z&\colon x^2+y^2 \to 2. \end{aligned} \tag{21.430}$$

```
----------------------------------------------------------------------
| 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(z-x**2)).integrate(x,0,sqrt(z)).integrate(z,0,2)
(x, y, z) |--> 8/15*sqrt(2)
```

Écrivons le détail de cette dernière intégrale :

$$\begin{aligned} I &= \int_0^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} dy \int_{x^2+y^2}^{2} x dz \\ &= \int_0^{\sqrt{2}} dx \int_0^{\sqrt{2-x^2}} x(2-x^2-y^2) dy \\ &= \int_0^{\sqrt{2}} dx\, x \left[(2-x^2)y - \frac{y^3}{3} \right]_0^{\sqrt{2-x^2}} \\ &= \int_0^{\sqrt{2}} \frac{2}{3} x(2-x^2)^{3/2} dx. \end{aligned} \tag{21.431}$$

Ici nous effectuons le changement de variable $u = x^2$, $du = 2xdx$. Ne pas oublier de changer les bornes de l'intégrale :

$$I = \frac{1}{3} \int_0^2 (2-u)^{3/2} du. \tag{21.432}$$

Le changement de variable $t = 2 - u$, $dt = -du$ fait venir (attention aux bornes ! !)

$$I = -\frac{1}{3} \int_2^0 t^{3/2} dt = \frac{1}{3} \left[\frac{t^{5/2}}{5/2} \right]_0^2 = \frac{8}{15}\sqrt{2}. \tag{21.433}$$

△

21.19.1 Volume

Parmi le nombreuses interprétations géométriques de l'intégrale triple, notons celle-ci :

Proposition 21.120.
Soit $D \subset \mathbb{R}^3$. Le volume de D est donné par

$$Vol(D) = \int_D dxdydz. \tag{21.434}$$

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.

Exemple 21.121
Calculons le volume de la sphère de rayon R. Le domaine de variation des variables x, y et z pour la sphère est

$$\begin{aligned} x&: -R \to R \\ y&: -\sqrt{R^2 - x^2} \to \sqrt{R^2 - x^2} \\ z&: -\sqrt{R^2 - x^2 - y^2} \to \sqrt{R^2 - x^2 - y^2}. \end{aligned} \tag{21.435}$$

Par conséquent nous devons calculer l'intégrale

$$V = \int_{-R}^{R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} dy \int_{-\sqrt{R^2-x^2-y^2}}^{\sqrt{R^2-x^2-y^2}} dz. \tag{21.436}$$

La première intégrale est simple :

$$V = 2\int_{-R}^{R} dx \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \sqrt{R^2 - x^2 - y^2}dy. \tag{21.437}$$

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

$$I = \int_{-\sqrt{a}}^{\sqrt{a}} \sqrt{a - y^2}dy, \tag{21.438}$$

ce qui est la surface du demi-disque de rayon $\sqrt{a}$. Nous avons donc

$$I = \frac{\pi a}{2} = \frac{\pi}{2}(R^2 - x^2), \tag{21.439}$$

et

$$V = 2\int_{-R}^{R} \frac{\pi}{2}(R^2 - x^2)dx = \pi\left[R^2 x - \frac{x^3}{3}\right]_{-R}^{R} = \frac{4}{3}\pi R^3. \tag{21.440}$$

△

Exemple 21.122
Nous pouvons calculer le volume de la sphère en utilisant les coordonnées sphériques. Les bornes des variables pour la sphère de rayon R sont

$$\begin{aligned} \rho&: 0 \to R \\ \theta&: 0 \to \pi \\ \varphi&: 0 \to 2\pi. \end{aligned} \tag{21.441}$$

En n'oubliant pas le jacobien $\rho^2 \sin(\theta)$, l'intégrale à calculer est

$$V = \int_0^R d\rho \int_0^{2\pi} d\varphi \int_0^{\pi} \rho^2 \sin(\theta)d\theta \tag{21.442}$$

L'intégrale sur φ fait juste une multiplication par 2π. Celle sur ρ vaut

$$\int_0^R \rho^2 d\rho = \frac{R^3}{3}. \tag{21.443}$$

L'intégrale sur θ donne

$$\int_0^\pi \sin(\theta)d\theta = [-\cos(\theta)]_0^\pi = 2. \tag{21.444}$$

Le tout fait par conséquent

$$V = \frac{4}{3}\pi R^3. \tag{21.445}$$

Sans contestes, le passage aux coordonnées sphériques a considérablement simplifié le calcul par rapport à celui de l'exemple 21.121. △

21.20 Un petit peu plus formel

21.20.1 Intégration sur un domaine non rectangulaire

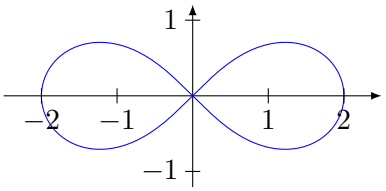


FIGURE 21.12 – Intégrer sur des domaines plus complexes.

La méthode de Fubini ne fonctionne plus sur un domaine non rectangulaire tel que celui de la figure 21.12. Nous allons donc utiliser une astuce. Considérons le domaine

$$E = \{(x, y) \in \mathbb{R}^2 \text{ tel que } a < x < b \text{ et } \alpha(x) < y < \beta(x)\} \tag{21.446}$$

représenté sur la figure 21.12. Nous considérons la fonction

$$\tilde{f}(x, y) = \begin{cases} f(x, y) & \text{si } (x, y) \in E \\ 0 & \text{sinon.} \end{cases} \tag{21.447}$$

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

$$\int_E f = \int_E \tilde{f} = \int_{\text{rectangle}} \tilde{f} = \int_a^b \left(\int_{\alpha(x)}^{\beta(x)} f(x, y)dy \right) dx. \tag{21.448}$$

Dans le cas de l'intégrale de $f(x, y) = x^2 + y^2$ sur le triangle de la figure 21.9, nous avons

$$\int_{\text{triangle}} (x^2 + y^2)dxdy = \int_0^2 \left(\int_0^y (x^2 + y^2)dx \right) dy. \tag{21.449}$$

Remarque 21.123.
Le nombre $\int_D f(x, y)dxdy$ ne dépend pas du choix du rectangle englobant D.

En pratique, nous calculons l'intégrale en utilisant une extension du théorème de Fubini :

Théorème 21.124.
Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction continue où D est un domaine de type vertical ou horizontal.

(1) Si D est vertical, alors

$$\int_D f = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dy \right] dx. \tag{21.450}$$

(2) Si D est horizontal, alors

$$\int_D f = \int_c^d \left[\int_{\psi_1(y)}^{\psi_2(y)} f(x,y)dx \right] dy. \tag{21.451}$$

21.20.2 Changement de variables

Le théorème du changement de variable est le suivant.

Théorème 21.125.
Soit $g\colon A \to B$ un difféomorphisme. Soient $F \subset B$ un ensemble mesurable et borné et $f\colon F \to \mathbb{R}$ une fonction bornée et intégrable. Supposons que $g^{-1}(F)$ soit borné et que Jg soit borné sur $g^{-1}(F)$. Alors

$$\int_F f(x)dy = \int_{g^{-1}(F)} f\big(g(x)\big)|Jg(x)|dx \tag{21.452}$$

Pour rappel, Jg est le déterminant de la matrice jacobienne (aucun lien de parenté) donnée par

$$Jg = \det \begin{pmatrix} \partial_x g_1 & \partial_y g_1 \\ \partial_x g_2 & \partial_t g_2 \end{pmatrix}. \tag{21.453}$$

Un **difféomorphisme** est une application $g\colon A \to B$ telle que g et $g^{-1}\colon B \to A$ soient de classe C^1.

21.20.2.1 Coordonnées polaires

Les coordonnées polaires sont données par le difféomorphisme

$$\begin{aligned} g\colon]0,\infty[\times]0,2\pi[&\to \mathbb{R}^2\backslash D \\ (r,\theta) &\mapsto \big(r\cos(\theta), r\sin(\theta)\big) \end{aligned} \tag{21.454}$$

où D est la demi-droite $y = 0$, $x \geqslant 0$. 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

$$Jg = \det \begin{pmatrix} \partial_r x & \partial_\theta x \\ \partial_r y & \partial_\theta y \end{pmatrix} = \det \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} = r. \tag{21.455}$$

21.20.2.2 Coordonnées sphériques

Les coordonnées sphériques sont données par

$$\left\{ \begin{array}{lll} x = r\cos\theta\sin\varphi & & r \in]0,\infty[\\ y = r\sin\theta\sin\varphi & \text{avec} & \theta \in]0,2\pi[\\ z = r\cos\varphi & & \phi \in]0,\pi[. \end{array} \right. \tag{21.456}$$

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

$$\int_0^R dr \int_0^{2\pi} d\theta \int_0^{\pi} r^2\sin(\phi)d\phi = 4\pi R = \frac{4}{3}\pi R^3. \tag{21.457}$$

Ici, la valeur absolue n'est pas importante parce que lorsque $\phi \in]0,\pi,[$, le sinus de ϕ est positif.

Des petits malins pourraient remarquer que le changement de variable (21.456) est encore une paramétrisation de $\mathbb{R}^3$ si on intervertit le domaine des angles :

$$\begin{aligned}\theta&\colon 0\to\pi\\ \phi&\colon 0\to 2\pi,\end{aligned}\tag{21.458}$$

alors nous paramétrons encore parfaitement bien la sphère, mais hélas

$$\int_0^R dr\int_0^\pi d\theta\int_0^{2\pi} r^2\sin(\phi)d\phi=0.\tag{21.459}$$

Pourquoi ces « nouvelles » coordonnées sphériques sont-elles mauvaises ? Il y a que quand l'angle ϕ parcours $]0,2\pi[$, 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 ϕ entre 0 et π, puis $-r^2\sin(\phi)$ pour ϕ entre π et 2π. Donc l'intégrale (21.459) n'est pas correcte. Il faut la remplacer par

$$\int_0^R dr\int_0^\pi d\theta\int_0^\pi r^2\sin(\phi)d\phi-\int_0^R dr\int_0^\pi d\theta\int_\pi^{2\pi} r^2\sin(\phi)d\phi=\frac{4}{3}\pi R^3\tag{21.460}$$

21.21 Aire et primitive

Soient $f\colon\mathbb{R}\to\mathbb{R}$ une fonction continue et $x\in\mathbb{R}$. Pour chaque $x\in\mathbb{R}$, nous pouvons considérer le nombre $F(x)$ défini par

$$F(x)=\int_a^x f(t)dt.\tag{21.461}$$

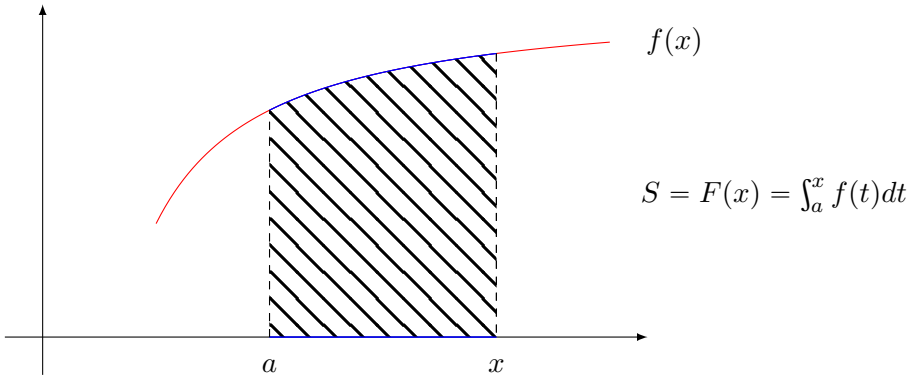


FIGURE 21.13 – Surface sous une courbe

La fonction F ainsi définie a deux importantes propriétés :

(1) C'est une primitive de f,

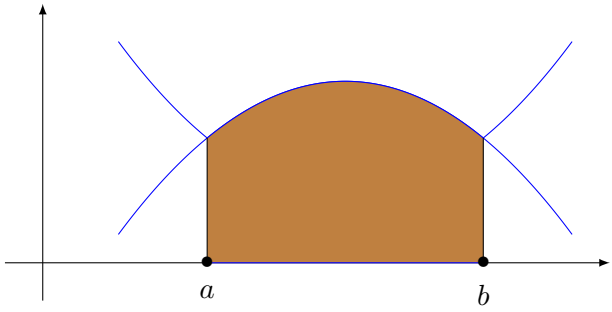
(2) Elle donne la surface en dessous de f entre les points a et x, voir la figure 21.13.

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 21.14.
La surface entre les deux fonctions $y_1(x)$ et $y_2(x)$ se calcule comme suit.

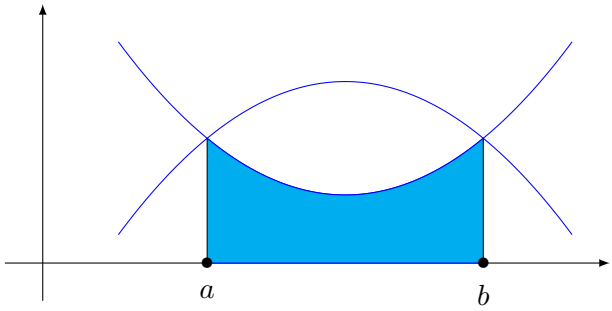
(1) On calcule les intersections entre $y1$ et y_2. Notons a et b les ordonnées obtenues.

(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

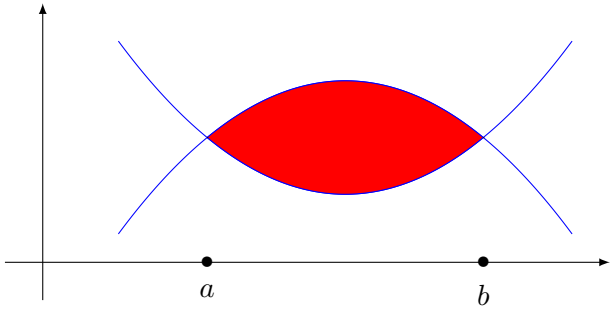
$$S=\int_a^b y_1-\int_a^b y_1.\tag{21.462}$$



(a) La grande surface.



(b) La surface à soustraire.



(c) La surface entre les deux fonctions.

FIGURE 21.14 – Le calcul de la surface comprise entre deux fonctions.

21.21.1 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

$$l(x_0, x_1) = \int_{x_0}^{x_1} \sqrt{1 + y'(t)^2} dt. \tag{21.463}$$

Lorsque la courbe est donnée sous forme paramétrique

$$\begin{cases} x = x(t) & \text{(21.464a)} \\ y = y(t), & \text{(21.464b)} \end{cases}$$

alors la formule devient

$$l(t_1, t_2) = \int_{t_1}^{t_2} \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} dt, \tag{21.465}$$

où $\dot{x}(t) = x'(t)$.

21.21.2 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

$$S = 2\pi \int_a^b \sqrt{1 + f'(x)^2} f(x) dx. \tag{21.466}$$

21.22 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 21.15.

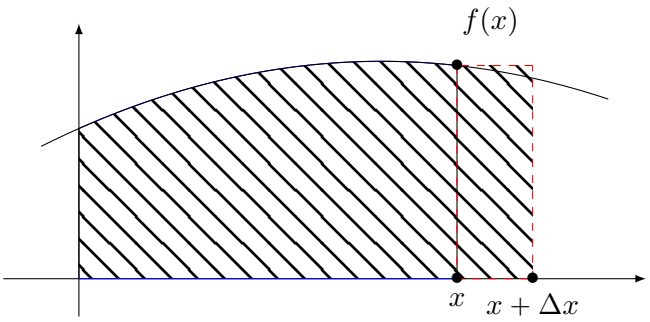


FIGURE 21.15 – L'aire en dessous d'une courbe. Le rectangle rouge d'aire $f(x)\Delta x$ approxime de combien la surface augmente lorsqu'on passe de x à $x + \Delta x$.

Si la fonction f est continue et que Δx 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 Δx est très petit,

$$S_f(x + \Delta x) - S_f(x) = f(x)\Delta x, \tag{21.467}$$

ou encore

$$f(x) = \frac{S_f(x + \Delta x) - S_f(x)}{\Delta x}. \tag{21.468}$$

Nous formalisons la notion de « lorsque Δx est très petit » par une limite :

$$f(x) = \lim_{\Delta x \to 0} \frac{S_f(x + \Delta x) - S_f(x)}{\Delta x}. \tag{21.469}$$

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.

21.23 Propriétés des intégrales

Lemme 21.126.
Pour toute fonction $u\colon [a,b] \to \mathbb{R}^n$, nous avons

$$\| \int_a^b u(t)dt \| \leqslant \int_a^b \|u(t)\| dt \tag{21.470}$$

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 11.293, il existe un $\xi \in \mathbb{R}^n$ de norme 1 tel que

$$\| \int_a^b u(t)dt \| = \xi \cdot \int_a^b u(t)dt = \int_a^b u(t) \cdot \xi dt \leqslant \int_a^b \|u(t)\|\|\xi\| = \int_a^b \|u(t)\|dt. \tag{21.471}$$

□

Proposition 21.127 (Relations de Chasles)**.**
Soit f une fonction continue sur l'intervalle I. Si $a, b, c \in I$ nous avons

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx. \tag{21.472}$$

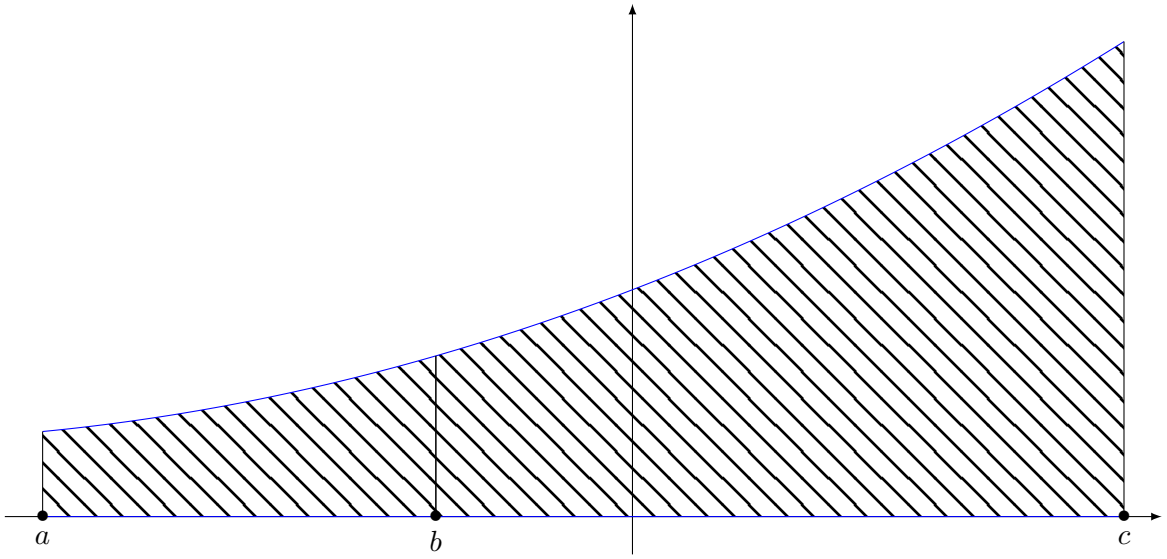


FIGURE 21.16 – Illustration pour les relations de Chasles.

Sur la figure 21.16, la surface de a à c est évidemment égale à la somme des surfaces de a à b et de b à c.

Corollaire 21.128.

$$\int_a^b f(x)dx = -\int_b^a f(x)dx. \tag{21.473}$$

Proposition 21.129 (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

$$\int_a^b \big(f(x) + g(x)\big)dx = \int_a^b f(x)dx + \int_a^b g(x)dx, \tag{21.474}$$

et

$$\int_a^b \lambda f(x)dx = \lambda \int_a^b f(x)dx. \tag{21.475}$$

Proposition 21.130 (L'intégrale est monotone)**.**
Soient $a, b \in I$ avec $a < b$. Si $f \geqslant g$ sur $[a, b]$ alors

$$\int_a^b f(x)dx \geqslant \int_a^b g(x)dx. \tag{21.476}$$

Corollaire 21.131 (Positivité)**.**
Si $a < b$ et $f \geqslant 0$ sur $[a, b]$ alors

$$\int_a^b f(x)dx \geqslant 0. \tag{21.477}$$

Ce résultat n'est qu'une application de la proposition 21.130 car il consiste à prendre comme fonction g la fonction nulle.

21.24 Techniques d'intégration

Par le théorème 15.233, la calcul d'une intégrale consiste essentiellement à trouver une primitive de la fonction à intégrer. Il est donc indispensable de bien connaitre les dérivées des fonctions usuelles.

Voici un tableau des primitives à connaitre.

Fonction $f(x)$	Primitive $\int f(x)\,dx$	Ensemble de définition de f	Remarques
x^α	$\frac{x^{\alpha+1}}{\alpha+1}+C$	dépend de α	$\alpha\in\mathbb{R}\backslash\{-1\}$
$\frac{1}{x}$	$\ln\big(\lvert x\rvert\big)+C$	$x\neq 0$	
$\frac{1}{1+x^2}$	$\arctan(x)+C$	$\mathbb{R}$	
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)+C$	$]-1,1[$	
$\frac{-1}{\sqrt{1-x^2}}$	$\arccos(x)+C$	$]-1,1[$	
e^x	e^x+C	$\mathbb{R}$	
$\sin(x)$	$-\cos(x)+C$	$\mathbb{R}$	
$\cos(x)$	$\sin(x)+C$	$\mathbb{R}$	
$1+\tan^2(x)$	$\tan(x)+C$	in 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 21.129. 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.

21.24.1 Intégration par parties

Proposition 21.132.
Si u et v sont deux fonctions dérivables de dérivées continues sur l'intervalle $[a,b]$ alors

$$\int_a^b u(x)v'(x)dx=\Big[u(x)v(x)\Big]_a^b-\int_a^b u'(x)v(x)dx. \tag{21.478}$$

Démonstration. Il s'agit d'utiliser à l'envers la formule de dérivation d'un produit :

$$uv'=(uv)'-u'v. \tag{21.479}$$

Les fonctions à gauche et à droite étant égales, elles ont même intégrale sur $[a,b]$ et par linéarité, voir proposition 21.129, on a :

$$\int_a^b u(x)v'(x)dx=\int_a^b (uv)'(x)-\int_a^b u'(x)v(x)dx. \tag{21.480}$$

La fonction uv est évidemment une primitive de $(uv)'$, de telle sorte que l'on puisse un peu simplifier cette expression :

$$\int_a^b u(x)v'(x)dx=\Big[u(x)v(x)\Big]_a^b-\int_a^b u'(x)v(x)dx, \tag{21.481}$$

ce qu'il fallait démontrer. □

Exemple 21.133
Un cas typique d'utilisation de l'intégrale par parties est le suivant. Soit à calculer

$$\int_0^\pi x\cos(x)dx. \tag{21.482}$$

Nous devons écrire $x\cos(x)$ comme un produit $u(x)v'(x)$. Il y a (au moins) deux moyens de le faire :

$$\begin{cases} u=x & \text{(21.483a)}\\ v'=\cos(x). & \text{(21.483b)}\end{cases}$$

ou

$$\begin{cases} u=\cos(x) & \text{(21.484a)}\\ v'=x. & \text{(21.484b)}\end{cases}$$

Nous allons choisir le premier [21]. Nous avons donc

$$\begin{aligned} u = x, \quad & v' = \cos(x) \\ u' = 1 \quad & v = \sin(x). \end{aligned} \tag{21.485}$$

En utilisant la formule d'intégration par parties,

$$\int_0^\pi x\cos(x)dx = \Big[x\sin(x)\Big]_0^\pi - \int_0^\pi 1\times\sin(x)dx = \pi\sin(\pi) - \Big[-\cos(x)\Big]_0^\pi = -2. \tag{21.486}$$

△

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)

$$\int uv' = uv - \int u'v. \tag{21.487}$$

Exemple 21.134
Nous reprenons l'exemple 21.133 en déterminant cette fois une primitive de $x\cos(x)$:

$$\int x\cos(x)dx = x\sin(x) - \int \sin(x)dx = x\sin(x) + \cos(x) + C, \qquad C\in\mathbb{R}. \tag{21.488}$$

Nous retrouvons le résultat numérique de l'exemple précédent en ajoutant les extrêmes d'intégration

$$\int_0^\pi x\cos(x)dx = \big[x\sin(x) + \cos(x)\big]_0^\pi = -2. \tag{21.489}$$

△

Remarque 21.135.
Lorsqu'on calcule des intégrales, il est bon de passer par la primitive (c'est-à-dire en suivant l'exemple 21.134 et non 21.133) parce qu'il est alors facile de vérifier le résultat en calculant la dérivée de la primitive trouvée.

Par exemple pour vérifier si (21.488) est correct, il suffit de dériver $x\sin(x)+\cos(x)$:

$$\big(x\sin(x)+\cos(x)\big)' = \sin(x) + x\cos(x) - \sin(x) = x\cos(x). \tag{21.490}$$

La fonction $x\sin(x)+\cos(x)$ est donc bien une primitive de $x\cos(x)$.

Exemple 21.136(Primitive du logarithme)
La primitive de la fonction logarithme définie en 16.58 nous offre un bon moment d'intégration par partie.

Trouver la primitive de la fonction $x\mapsto\ln(x)$. Pour calculer

$$\int \ln(x)dx \tag{21.491}$$

nous écrivons $\ln(x) = 1\times\ln(x)$ et nous posons $u'=1$ et $v=\ln(x)$, c'est-à-dire

$$\begin{aligned} u' = 1 \quad & v = \ln(x) \\ u = x \quad & v' = \frac{1}{x}. \end{aligned} \tag{21.492}$$

La formule d'intégration par parties (21.478) donne donc

$$\int \ln(x) = x\ln(x) - \int x\times\frac{1}{x} = x\ln(x) - \int 1 = x\ln(x) - x + C, \qquad C\in\mathbb{R}. \tag{21.493}$$

21. Mais nous conseillons vivement au lecteur d'essayer le deuxième pour se rendre compte qu'il ne fonctionne pas.

Il est facile de vérifier par un petit calcul que

$$\big(x\ln(x)-x\big)' = \ln(x). \tag{21.494}$$

△

21.24.2 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 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.

Proposition 21.137.
Soient I et J des intervalles de $\mathbb{R}$, $u\colon I\to J$ une fonction qui est dérivable de dérivée continue et $f\colon J\to\mathbb{R}$ une fonction admettant une primitive F. Alors la fonction

$$x\mapsto F\big(u(x)\big) \tag{21.495}$$

est une primitive de

$$f\big(u(x)\big)u'(x). \tag{21.496}$$

Démonstration. Cela est une utilisation immédiate de la formule de dérivée des fonctions composées. □

Exemple 21.138
Soit à calculer

$$\int x\sqrt{1-x^2}dx. \tag{21.497}$$

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 (21.496) 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,

$$\int -2x\sqrt{1-x^2}\,dx = \frac{2(1-x^2)^{3/2}}{3}+C, \tag{21.498}$$

et en divisant par -2 nous trouvons la primitive demandée :

$$\int x\sqrt{1-x^2}\,dx = -\frac{(1-x^2)^{3/2}}{3}+C. \tag{21.499}$$

△

L'exemple suivant donne une façon plus économe de retenir la méthode du changement de variables.

Exemple 21.139
Soit à calculer

$$\int \cos(x)e^{\sin(x)}dx. \tag{21.500}$$

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

$$du = \cos(x)dx. \tag{21.501}$$

Nous avons donc les substitutions suivantes à faire :

$$\sin(x) = u \tag{21.502a}$$
$$du = \cos(x)dx \tag{21.502b}$$
$$dx = \frac{du}{\cos(x)}. \tag{21.502c}$$

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

$$\int \cos(x)e^{\sin(x)}dx = \int e^u du = e^u + C, \qquad \text{où } u = \sin(x). \tag{21.503}$$

Attention : la réponse doit être impérativement donnée en termes de x et non de u. Nous écrivons donc

$$\int \cos(x)e^{\sin(x)} = e^{\sin(x)} + C. \tag{21.504}$$

△

21.24.3 Changement de variables – pour calculer des intégrales

Le théorème 15.233 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 corollaire 15.233 ;

— écrire une intégrale pour la nouvelle variable $u = u(x)$ sur l'intervalle entre $u(a)$ et $u(b)$.

Nous allons voir ce deux méthodes dans des exemples.

Exemple 21.140
Soit à calculer

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx. \tag{21.505}$$

Les primitives $\int x\sqrt{1-x^2}dx$ ont été trouvé dans l'exemple 21.139. Une primitive est

$$F(x) = \int x\sqrt{1-x^2}dx = -\frac{(1-x^2)^{3/2}}{3}. \tag{21.506}$$

Nous pouvons maintenant calculer l'intégrale de $x\sqrt{1-x^2}$ sur l'intervalle $[1/3, 1/2]$ par la définition

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx = F\left(\frac{1}{2}\right) - F\left(\frac{1}{3}\right) = -\frac{\sqrt{3}}{8} + \frac{16\sqrt{2}}{81}. \tag{21.507}$$

△

Remarque 21.141.
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 21.142
Calculons à nouveau

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx. \tag{21.508}$$

Cette fois nous allons toucher à l'intervalle d'intégration en même temps que faire le changement de variables. Nous savons déjà les substitutions

$$\begin{cases} u = 1 - x^2 & (21.509a)\\ du = -2xdx & (21.509b)\\ dx = \dfrac{du}{-2x}. & (21.509c) \end{cases}$$

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 :

$$\begin{cases} x = 1/3 \to u = 8/9 & (21.510a)\\ x = 1/2 \to u = 3/4 & (21.510b) \end{cases}$$

Le calcul est alors

$$\int_{1/3}^{1/2} x\sqrt{1-x^2}dx = -\frac{1}{2}\int_{8/9}^{3/4}\sqrt{u}du = -\frac{1}{2}\left[\frac{u^{3/2}}{3/2}\right]_{8/9}^{3/4} = -\frac{\sqrt{3}}{8} + \frac{16\sqrt{2}}{81}. \tag{21.511}$$

Attention : la dernière égalité n'est pas immédiate ; elle demande quelques calculs et une bonne utilisation des règles de puissances. △

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.

Exemple 21.143
À calculer :

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx. \tag{21.512}$$

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 [22]. Les substitutions à faire dans l'intégrale sont :

$$\begin{cases} x = \sin(\theta) & (21.513a)\\ dx = \cos(\theta)d\theta, & (21.513b) \end{cases}$$

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

$$\int_{1/2}^{\sqrt{3}/2} \sqrt{1-x^2}dx = \int_{\pi/6}^{\pi/3} \sqrt{1-\sin^2(\theta)}\cos(\theta)dt. \tag{21.514}$$

22. 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.

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

$$\int_{\pi/6}^{\pi/3}\cos^2(\theta)d\theta. \tag{21.515}$$

Pour celle-là, il faut utiliser une formule de trigonométrie [23] :

$$\cos^2(\theta)=\frac{1+\cos(2\theta)}{2}. \tag{21.516}$$

Donc

$$\int_{\pi/6}^{\pi/3}\cos^2(\theta)d\theta=\int_{\pi/6}^{\pi/3}\frac{1+\cos(2\theta)}{2}d\theta=\left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3}+\int_{\pi/6}^{\pi/3}\frac{\cos(2\theta)}{2}d\theta, \tag{21.517}$$

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

$$\int_{\pi/6}^{\pi/3}\cos^2(\theta)d\theta=\left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3}+\int_{\pi/3}^{2\pi/3}\frac{\cos(t)}{4}dt=\left[\frac{\theta}{2}\right]_{\pi/6}^{\pi/3}=\frac{\pi}{6}-\frac{\pi}{12}=\frac{\pi}{12}, \tag{21.518}$$

parce que $\sin\left(\frac{2\pi}{3}\right)=\sin\left(\frac{\pi}{3}\right)$. Au final,

$$\int_{1/2}^{\sqrt{3}/2}\sqrt{1-x^2}dx=\frac{\pi}{12}. \tag{21.519}$$

△

21.24.4 Intégrations des fractions rationnelles réduites

Définition 21.144.
*Une **fraction rationnelle** est un quotient de deux polynômes à coefficients réels ou complexes.*

Par exemple

$$\frac{x^5+7x^4-\frac{x^3}{2}+x}{x^2-1} \tag{21.520}$$

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 :

$$\frac{\alpha}{(x-a)^m}, \qquad \alpha,a\in\mathbb{R},m\in\mathbb{N} \tag{21.521a}$$

$$\frac{\alpha x+\beta}{(x^2+ax+b)^m}; \qquad \alpha,\beta,a,b\in\mathbb{R},m\in\mathbb{N},a^2-4b<0. \tag{21.521b}$$

Nous allons nous contenter de donner un exemple de chaque type.

(1) En ce qui concerne le cas (21.521a) avec $m=1$, nous avons par exemple

$$\int\frac{1}{x-3}dx=\ln\left(|x-3|\right)+C. \tag{21.522}$$

Si vous voulez en être tout à fait sûr, effectuez d'abord le changement de variables $u=x-3$ qui donne $dx=du$.

23. En fait, il y a moyen de terminer le calcul en intégrant deux fois par parties, mais c'est plus compliqué.

(2) En ce qui concerne le cas (21.521a) avec $m \neq 1$, nous avons par exemple

$$\int \frac{1}{(x-1)^4} dx = -\frac{1}{3(x-1)^3} + C. \tag{21.523}$$

Encore une fois, pour s'en convaincre, utiliser le changement de variables $u = x-1$, $dx = du$:

$$\int \frac{1}{(x-1)^4} dx = \int \frac{1}{u^4} du = \int u^{-4} du = -\frac{u^{-3}}{3} + C = -\frac{1}{3}\frac{1}{(x-1)^3} + C. \tag{21.524}$$

(3) En ce qui concerne le cas (21.521b) avec $\alpha \neq 0$, nous avons par exemple

$$\int \frac{x}{x^2+4} dx = \frac{1}{2}\ln(x^2+4) + C. \tag{21.525}$$

Pour ce faire, il faut faire le changement de variables $u = x^2+4$, $du = 2xdx$, $dx = \frac{du}{2x}$ qui donne

$$\int \frac{x}{x^2+4} dx = \frac{1}{2}\int \frac{du}{u} = \frac{1}{2}\ln(|u|) + C = \frac{1}{2}\ln(|x^2+4|) + C. \tag{21.526}$$

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 (21.521b) avec $\alpha = 0$, nous avons par exemple

$$\int \frac{dx}{x^2+4} = \frac{1}{4}\int \frac{dx}{(\frac{x}{2})^2+1} = \frac{1}{2}\arctan(\frac{x}{2}) + C. \tag{21.527}$$

où nous avons utilisé la primitive $\int \frac{dx}{x^2+1} dx = \arctan(x)$ du tableau de la page 1414. Pour vous en convaincre vous pouvez faire la dernière étape avec le changement de variables $u = x/2$, $dx = 2du$.

21.24.5 Quelques formules à connaitre

À retenir 21.145

$$\int (\alpha f(x) + \beta g(x))\, dx = \alpha \int f(x)\, dx + \beta \int g(x)\, dx. \tag{21.528a}$$

$$\int f(x)g'(x)\, dx = f(x)g(x) - \int f'(x)g(x)\, dx. \tag{21.528b}$$

$$\int f'(u(x))u'(x)\, dx = \int f(t)\, dt, \qquad \text{avec } t = u(x). \tag{21.528c}$$

$$\int \frac{f'(x)}{f(x)}\, dx = \log|f(x)| + C, \qquad \text{c'est un cas particulier de la formule précédente.} \tag{21.528d}$$

21.24.6 Approximation de $\ln(2)$

Théorème 21.146.
Soit une fonction $f\colon [a,b] \to \mathbb{R}$ de classe C^{n+1}. Alors pour tout $N \leqslant n$ nous avons

$$f(b) = f(a) + \sum_{k=1}^{N} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{1}{N!}\int_a^b (b-t)^N f^{(N+1)}(t) dt. \tag{21.529}$$

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 (21.529) nous avons

$$f(a) + f'(a)(b-a) \int_a^b (b-t)f''(t)dt. \tag{21.530}$$

Nous évaluons l'intégrale à part en faisant une intégration par parties [24]. Il s'agit de poser

$$u = b - t \tag{21.531a}$$
$$v' = f'', \tag{21.531b}$$

de déduire

$$u' = -1 \tag{21.532a}$$
$$v = f' \tag{21.532b}$$

et d'écrire

$$\int_a^b (b-t)f''(t)dt = \big[(b-t)f'(t)\big]_a^b - \int_a^b (-)f'(t)dt \tag{21.533a}$$
$$= -(b-a)f'(a) + \int_a^b f'(t)dt \tag{21.533b}$$
$$= -(b-a)f'(a) + f(b) - f(a). \tag{21.533c}$$

Dans le calcul nous avons utilisé le théorème fondamental du calcul intégral 15.233. En remettant ça dans (21.530) nous trouvons $f(b)$ comme il se doit.

En ce qui concerne la récurrence, nous devons calculer

$$f(a) + \sum_{k=1}^{N+1} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{1}{(N+1)!} \int_a^b (b-t)^{N+1} f^{(N+2)}(t)dt. \tag{21.534}$$

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

$$\int_a^b (b-t)^{N+1} f^{(N+2)}(t)dt = -(b-a)^{N+1} f^{(N+1)}(a) + (N+1) \int_a^b (b-t)^N f^{(N+1)}(t)dt. \tag{21.535}$$

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 (21.534) se réduit bien à $f(b)$. □

Cette formule avec reste intégral sert par exemple à prouver un encadrement pour $\ln(2)$, voir la proposition 21.147.

Proposition 21.147 (Approximation de $\ln(2)$[358])**.**
Pour tout n nous avons

$$\left| \ln(2) - \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k} \right| \leqslant \frac{1}{n+1}. \tag{21.536}$$

24. Proposition 21.132.

Démonstration. Nous écrivons la formule de Taylor avec reste intégral du théorème 21.146 pour la fonction $f = \ln$ et pour $a = 1$ et $b = x$. Cela donne :

$$\ln(x) = \ln(1) + \sum_{k=1}^{N} \frac{\ln^{(k)}(1)}{k!}(x-1)^k + \frac{1}{N!}(x-t)^N \ln^{(N+1)}(t)dt. \tag{21.537}$$

Sachant que la dérivée du logarithme [25] est $1/x$ et faisant une petite récurrence, pour $k \geqslant 1$ nous avons

$$\ln^{(k)}(x) = \frac{(-1)^{k+1}(k-1)!}{x^k}. \tag{21.538}$$

En remplaçant,

$$\ln(x) = \sum_{k=1}^{N} \frac{(-1)^{k+1}}{k}(x-1)^k + \int_1^x \frac{(-1)^N(x-t)^N}{t^{N+1}}dt. \tag{21.539}$$

C'est le moment de poser $x = 2$ et de faire les simplifications qui s'imposent,

$$\ln(2) = \sum_{k=1}^{N} \frac{(-1)^{k+1}}{k} + \int_1^2 \frac{(-1)^N(2-t)^N}{t^{N+1}}dt. \tag{21.540}$$

Nous déplaçons la somme à gauche, et nous prenons la valeur absolue des deux côtés :

$$|\ln(2) - \sum_{k=1}^{N} \frac{(-1)^{k+1}}{k}| = |\int_1^2 \frac{(-1)^N(2-t)^N}{t^{N+1}}dt| \tag{21.541a}$$

$$\leqslant \int_1^2 \int_1^2 \frac{(2-t)^N}{t^{N+1}}dt \tag{21.541b}$$

$$\leqslant \int_1^2 (2-t)^N. \tag{21.541c}$$

Justifications :

— Pour 21.541b. Majoration en rentrant la valeur absolue dans l'intégrale, suppression de $(-1)^N$, et le fait que pour $t \in [1,2]$, $2 - t \geqslant 0$.

— Pour 21.541c. Majoration en supprimant purement et simplement le dénominateur $t^N + 1 \geqslant 1$.

Ais-je vraiment besoin de vous dire que la dernière intégrale se calcule en posant le changement de variables [26] $u = 2 - t$? Le résultat est que

$$\int_1^2 (2-t)^N dt = \frac{1}{N+1}. \tag{21.542}$$

□

Exemple 21.148([1])
La convergence de l'encadrement (21.147) n'est pas terrible. Pour avoir une erreur de $\frac{1}{10}$, il faut

$$\frac{1}{10} = \frac{1}{n+1}, \tag{21.543}$$

ce qui demande $n = 9$. Ça reste jouable, même pour les jeunes d'aujourd'hui. Écrivons 9 termes :

$$|\ln(2) - 1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \frac{1}{5} + \frac{1}{6} - \frac{1}{7} + \frac{1}{8} - \frac{1}{9}| \leqslant \frac{1}{10}. \tag{21.544}$$

En calculant [27],

$$|\ln(2) - \frac{1879}{2520}| \leqslant \frac{1}{10}. \tag{21.545}$$

25. Voir la proposition 16.64.
26. Proposition 21.137.
27. Moi j'ai utilisé Sage, mais si tu es au tableau, débrouilles-toi.

Voici donc un bel encadrement

$$\frac{1879}{2520}-\frac{1}{10}\leqslant \ln(2)\leqslant \frac{1879}{2520}+\frac{1}{10}. \tag{21.546}$$

Pour avoir quelque chose avec des virgules, d'abord un peu de calcul mental donne

$$\frac{1879}{2520}\simeq 0.745634920634921. \tag{21.547}$$

Donc en majorant et minorant, disons, la troisième décimale [28], on n'est pas moins précis que le $\frac{1}{10}$. On a

$$0.744-\frac{1}{10}\leqslant \ln(2)\leqslant 0.746+\frac{1}{10}. \tag{21.548}$$

Bref, on retient l'approximation

$$0.644\leqslant \ln(2)\leqslant 0.846. \tag{21.549}$$

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. △

Proposition 21.149 (Formule de Taylor avec reste intégral[359, 223])**.**
Soient X et Y des espaces normés et un ouvert $\mathcal{O}\subset X$. Si $f\in C^m(\mathcal{O},Y)$ et si $[p,x]\subset\mathcal{O}$ alors

$$\begin{aligned} f(x)=f(p)&+\sum_{k=1}^{m-1}\frac{1}{k!}(d^kf)_p(x-p)^k\\ &+\frac{1}{(m-1)!}\int_0^1(1-t)^{m-1}(d^mf)_{p+t(x-p)}(x-p)^m\,dt \end{aligned} \tag{21.550}$$

où $\omega_p u^k$ signifie $\omega_p(u,\dots,u)$ lorsque $\omega\in\Omega^k$.

Notez que l'intégrale n'est pas une intégrale faisant intervenir les espaces X ou Y. Elle est une simple intégrale d'une fonction $\mathbb{R}\to\mathbb{R}$, comme définie par la mesure de Lebesgue de la définition 15.127.

Comme expliqué dans l'exemple 13.270, toute ces applications de différentielles se réduisent à des termes de la forme

$$f^{(k)}(p)(x-p)^k \tag{21.551}$$

dans le cas d'une fonction $\mathbb{R}\to\mathbb{R}$.

21.24.7 Lemme de Morse

Lemme 21.150 (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 $df_0=0$ et que d^2f_0 est non dégénérée [29] et de signature $(p,n-p)$. Alors il existe un C^1-difféomorphisme φ entre deux voisinages de 0 dans $\mathbb{R}^n$ tel que

(1) $\varphi(0)=0$,

(2) si $\varphi(x)=u$ alors

$$f(x)-f(0)=u_1^2+\cdots+u_p^2-u_{p+1}^2-\dots-u_n^2. \tag{21.552}$$

Une autre façon de dire est qu'il existe un C^1-difféomorphisme local ψ tel que

$$(f\circ\psi)(x)-f(0)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\dots-x_n^2. \tag{21.553}$$

28. Notez ici que nous utilisons le fait que la division euclidienne, elle, donne un encadrement pour les fractions. Pensez-y.

29. En tant qu'application bilinéaire.

Démonstration. Nous allons noter Hf la matrice hessienne de f, c'est-à-dire $Hf_a = d^2 f_a \in \mathcal{L}^{(2)}(\mathbb{R}^n, \mathbb{R})$. Écrivons la formule de Taylor avec reste intégral (proposition 21.149 avec $p = 0$ et $m = 2$) :

$$f(x) - f(0) = \underbrace{df_0(x)}_{=0} + \int_0^1 (1-t) \underbrace{d^2 f_{tx}(x,x)}_{x^t (Hf)_{tx} x = \langle Hf_{tx} x, x\rangle} dt = x^t Q(x) x \tag{21.554}$$

avec

$$Q(x) = \int_0^1 (1-t)(Hf)_{tx} dt \tag{21.555}$$

qui est une intégrale dans $\mathcal{L}^{(2)}(\mathbb{R}^n, \mathbb{R})$. Nous prouvons à présent que Q est de classe C^1 en utilisant le résultat de différentiabilité sous l'intégrale 18.26. Pour cela nous passons aux composantes (de la matrice) et nous considérons

$$\begin{aligned} h_{kl} &: U \times [0,1] \to \mathbb{R} \\ h_{kl}(x,t) &= (1-t) \frac{\partial^2 f}{\partial x_k \partial x_l}(tx). \end{aligned} \tag{21.556}$$

Étant donné que f est de classe C^3, la dérivée de h_{kl} par rapport à x_i ne pose pas de problèmes :

$$\frac{\partial h_{kl}}{\partial x_i} = t(t-1) \frac{\partial^3 f}{\partial x_i \partial x_k \partial x_l}(tx), \tag{21.557}$$

qui est encore continue à la fois en t et en x. La proposition 18.26 nous montre à présent que

$$Q_{kl}(x) = \int_0^1 (1-t) h_{kl}(tx) dt \tag{21.558}$$

est une fonction C^1. Étant donné que les composantes de Q sont C^1, la fonction Q est également C^1.

Nous avons $Q(0) = \frac{1}{2}(Hf)_0 \in S_n \cap \mathrm{GL}(n, \mathbb{R})$, d'abord parce que f est C^2 (et donc la matrice hessienne est symétrique), ensuite par hypothèse $d^2 f_0$ est non dégénérée.

À partir de là, le lemme 18.113 donne un voisinage V de $Q(0)$ dans S_n et une application ϕ de classe C^1

$$\phi : V \to \mathrm{GL}(n, \mathbb{R}) \tag{21.559}$$

telle que pour tout $A \in V$,

$$\phi(A)^t Q(0) \phi(A) = A. \tag{21.560}$$

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

$$Q(x) = M(x)^t Q(0) M(x). \tag{21.561}$$

Notons que l'application $\mathbb{M} : \mathbb{R} \to \mathrm{GL}(n, \mathbb{R})$ est de classe C^1 parce que Q et ϕ le sont.

Nous avons

$$f(x) - f(0) = x^t Q(x) x = x^t M(x)^t Q(0) M(x) x = y(x)^t Q(0) y(x) \tag{21.562}$$

où $y(x) = M(x)x = (\phi \circ Q)(x)x$ est encore une fonction de classe C^1 parce que la multiplication est une application C^∞.

D'un autre côté le théorème de Sylvester 4.110 nous donne une matrice inversible P telle que

$$Q(0) = P^t \begin{pmatrix} \mathbb{1}_p & \\ & -\mathbb{1}_{n-p} \end{pmatrix} P. \tag{21.563}$$

Et nous posons enfin $u = \varphi(x) = Py(x)$ qui est toujours de classe C^1 et qui donne

$$\begin{aligned} f(x) - f(0) &= y^tQ(0)y && (21.564a)\\ &= y^tP^t\begin{pmatrix}\mathbb{1} & \\ & -\mathbb{1}\end{pmatrix}Py && (21.564b)\\ &= u^t\begin{pmatrix}\mathbb{1} & \\ & -\mathbb{1}\end{pmatrix}u && (21.564c)\\ &= u_1^2 + \cdots + u_p^2 - u_{p+1}^2 - \ldots - u_n^2. && (21.564d)\end{aligned}$$

Nous devons maintenant montrer que, quitte à réduire son domaine à un ouvert plus petit, φ est un C^1-difféomorphisme. Dans la chaine qui donne φ, seule l'application

$$\begin{aligned} g\colon U \subset \mathbb{R}^n &\to \mathbb{R}^n\\ x &\mapsto M(x)x \end{aligned} \tag{21.565}$$

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 (13.482) :

$$dg_0(x) = \frac{d}{dt}\Big[g(tx)\Big]_{t=0} = \frac{d}{dt}\Big[tM(tx)x\Big]_{t=0} = M(0)x. \tag{21.566}$$

Note que nous avons utilisé la règle de Leibnitz pour la dérivée d'un produit, mais le second terme s'est annulé. Donc $dg_0 = M(0) \in \mathrm{GL}(n,\mathbb{R})$ et g est localement un C^1-difféomorphisme.

Il suffit de restreindre φ au domaine sur lequel g est un C^1-difféomorphisme pour que φ devienne lui-même un C^1-difféomorphisme.

□

Définition 21.151.
*Un point a est un **point critique** de la fonction différentiable f si $df_a = 0$.*

Corollaire 21.152 ([360])**.**
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 21.150, il existe un C^1-difféomorphisme ψ et un entier p tel que

$$(f \circ \psi)(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \ldots - x_n^2 + f(a) \tag{21.567}$$

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 :

$$\nabla(f \circ \psi)(x) = \sum_k \frac{\partial f}{\partial y_k}\big(g(x)\big)\nabla g_k(x). \tag{21.568}$$

Si $\psi(x)$ est une point critique de f, alors le membre de droite est le vecteur nul parce que tous les $\partial_k f\big(\psi(x)\big)$ 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. □

21.25 Constructions plus naïves de l'intégrale dans le cas réel

Les sections 15.2 et 15.8 ont donné une construction très complète de la mesure de Lebesgue, et nous avons définit la théorie de l'intégration sur un espace mesuré quelconque dans la définition 15.151.

Dans cette section nous allons donner différentes choses plus rapides qui servent souvent de définition dans les cours moins avancés.

21.25.1 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

$$B = \prod_{i=1}^{n} [a_i, b_i]; \tag{21.569}$$

le volume d'un tel pavé est défini par $\mathrm{Vol}(B) = \prod_i (b_i - a_i)$. Soit maintenant $A \subset \mathbb{R}^n$. La **mesure externe** de A est le nombre

$$m^*(A) = \inf\{ \sum_{B \in \mathcal{F}} \mathrm{Vol}(B) \text{ où } \mathcal{F} \text{ est un ensemble dénombrable de pavés dont l'union recouvre } A.\} \tag{21.570}$$

Définition 21.153.
*Nous disons que A est **mesurable** au sens de Lebesgue si pour tout ensemble $S \subset \mathbb{R}^n$ nous avons l'égalité*

$$m^*(S) = m^*(A \cap S) + m^*(S \backslash A). \tag{21.571}$$

Dans ce cas nous disons que la mesure de Lebesgue de A est $m(A) = m^*(A)$.

Proposition 21.154.
Deux fonctions continues égales presque partout pour la mesure de Lebesgue[30] *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. □

21.25.2 Pavés et subdivisions

Définition 21.155.
*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 $\mathbb{R}^p$ s'il s'écrit sous la forme*

$$R = \left\{ (x_1, \ldots, x_p) \in \mathbb{R}^p \,\middle|\, x_i \in \mathcal{I}_i, i = 1, \ldots, p \right\},$$

où $\mathcal{I}_i$ est un intervalle de $\mathbb{R}$ pour tout $i = 1, \ldots, p$.

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 $\mathbb{R}^p$.

30. Définition 21.153.

Définition 21.156.
Une partie A de $\mathbb{R}^p$ est dite ***pavable*** *s'il existe une famille finie de pavés bornés 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 21.17. Il existe beaucoup d'ensembles dans $\mathbb{R}^2$ qui ne sont pas pavables, par exemple les ellipses.

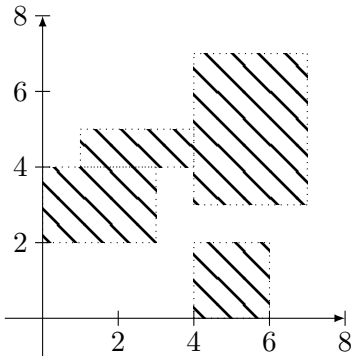


FIGURE 21.17 – Un ensemble pavable.

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.

Définition 21.157.
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 arrête 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 21.158
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 21.159
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 ∞. 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 22.3 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 de généralise au cas des pavés.

Définition 21.160.
Soir 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 $\mathcal{S} = \prod_{i=1}^{p} \mathcal{S}_i$,*

$$\mathcal{S} = \left\{ (Y_1, \dots, Y_p) \,\middle|\, Y_i = (y_{i,j})_{j=1}^{n_i} \in \mathcal{S}_i,\ i = 1, \dots, p \right\}.$$

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 21.18.

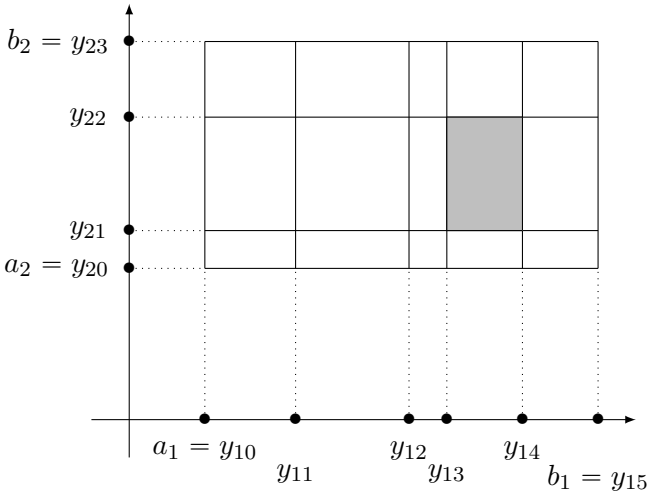


FIGURE 21.18 – Une cellule d'une subdivision d'un pavé de $\mathbb{R}^2$. La cellule grisée est $R_{(4,2)}$.

Définition 21.161.
Si σ est une subdivision d'un pavé R, un ***raffinement*** *de σ est une subdivision de R obtenue en fixant plus de points dans chaque intervalle.*

La subdivision σ de R détermine $n_1 n_2 \dots n_p$ pavés fermés de la forme

$$R_{(k_1,\dots,k_p)} = \{(x_1, \dots, x_p) \in \mathbb{R}^p \,|\, y_{i,k_i-1} \leqslant x_i \leqslant y_{i,k_i}\},$$

où k_i est dans $\{1, \dots, n_i\}$ et i dans $\{1, \dots, p\}$. On les appelles **cellules** de σ. On remarque que les cellules de σ sont toujours deux à deux disjointes (sauf au plus sur leurs bords).

Lemme 21.162.
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 21.162 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 21.163.
Une application $f : \mathbb{R}^p \to \mathbb{R}$ est dite ***application en escalier*** *sur $\mathbb{R}^m$ si*

- *f est une application bornée,*
- *il existe une subdivision σ de $\mathbb{R}^p$ telle que la restriction de f est une application constante sur toute cellule R_k de σ*

$$f_{|_{R_k}} = C_k, \qquad C_k \in \mathbb{R},$$

Une telle subdivision σ est dite ***associée*** *à f.*

Exemple 21.164
La fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y) = \begin{cases} 1 & \text{si } (x,y) \in [0,3] \times [-1,2], \\ 2 & \text{sinon.} \end{cases} \tag{21.572}$$

est une application en escalier. Exercice : donner une subdivision de $\mathbb{R}^2$ associée à cette fonction. △

Exemple 21.165
La fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y) = \begin{cases} \frac{1}{m^2+n^2}, & \text{si } (x,y) \in [m,m+1] \times [n,n+1], \quad m,\, n \in \mathbb{N}_0, \\ 0, & \text{sinon} \end{cases} \tag{21.573}$$

est une application en escalier. Observez que, dans ce cas, il n'existe pas une subdivision finie de $\mathbb{R}^2$ associée à f. △

Remarque 21.166.
Si la subdivision σ est associée à f alors tout raffinement de σ (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 σ_f et σ_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 $\sigma_f = (Y_1, \ldots, Y_p)$ et $\sigma_g = (Z_1, \ldots, Z_p)$, 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 $i = 1, \ldots, p$. 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 $\bar{\sigma} = (\bar{Y}_1, \ldots, \bar{Y}_p)$ de R est un raffinement de σ_f et de σ_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.

21.25.3 Intégrale d'une fonction en escalier

Définition 21.167.
*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 21.168.
*Une application en escalier f est dite **intégrable** si son support est compact.*

Soit f une application en escalier sur $\mathbb{R}^p$. Soit σ une subdivision de $\mathbb{R}^p$ associée à f et appelons R_k les cellules de σ, avec $k = (k_1, \ldots, k_p)$ dans $K = \{1, \ldots, n_1\} \times \{1, \ldots, n_2\} \times \ldots \times \{1, \ldots, n_p\}$. Alors

$$f_{|R_k} = C_k, \qquad C_k \in \mathbb{R}.$$

Définition 21.169.
*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 21.170.
Soit f une application en escalier intégrable sur $\mathbb{R}^p$. Soient σ_1 et σ_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.

21.25.4 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 21.171
Soit f la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x, y) = \begin{cases} x + 3y & \text{si } (x, y) \in [9, 10] \times]\pi, 5] \\ 0 & \text{sinon.} \end{cases} \tag{21.574}$$

Les intégrales partielles de f sont

$$F_1(y) = \int_9^{10} x + 3y\, dx = \left[\frac{x^2}{2} + 3xy\right]_{x=9}^{x=10} = \frac{19}{2} + 3y,$$

$$F_2(x) = \int_\pi^5 x + 3y\, dy = \left[xy + \frac{3y^2}{2}\right]_{y=\pi}^{y=5} = x(5 - \pi) + \frac{3}{2}(25 - \pi^2).$$

△

21.25.5 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 σ de R définie par les subdivisions

$$a = x_0 \leqslant x_1 \leqslant \ldots \leqslant x_m = b,$$
$$c = y_0 \leqslant y_1 \leqslant \ldots \leqslant y_n = d.$$

Les cellules de σ sont

$$R_{i,j} = [x_i, x_{i+1}] \times [y_j, y_{j+1}], \qquad i = 0, \dots, m-1, \quad j = 0, \dots, n-1.$$

La mesure de R est la somme des mesures des $R_{i,j}$

$$\begin{aligned} m(R) &= \sum_{(i,j) \in \{0,\dots,m-1\} \times \{0,\dots,n-1\}} m(R_{i,j}) = \\ &= \sum_{j=0}^{n-1} \sum_{i=0}^{m-1} (x_{i+1} - x_i) \cdot (y_{i+1} - y_i) = \\ &= \sum_{i=0}^{m-1} (x_{i+1} - x_i) \cdot \sum_{j=0}^{n-1} (y_{i+1} - y_i) = \\ &= (b - a) \cdot (d - c). \end{aligned} \tag{21.575}$$

Si f est constante sur chaque cellule de σ 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 $R_{i,j}$

$$\chi_{R_{i,j}}(x,y) = \begin{cases} 1, & \text{si } (x,y) \in R_{i,j}, \\ 0, & \text{sinon.} \end{cases} \tag{21.576}$$

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

$$\int_{\mathbb{R}^2} f\,dV = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,m(R_{i,j}) = \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,(x_{i+1}-x_i)\cdot(y_{j+1}-y_j). \tag{21.577}$$

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,

$$\begin{aligned} C_{i,j}(x_{i+1}-x_i)\cdot(y_{j+1}-y_j) &= \\ &= C_{i,j}\left(\int_{x_i}^{x_{i+1}} \chi_{[x_i,x_{i+1}]}(x)\,dx\right)\cdot\left(\int_{y_j}^{y_{j+1}} \chi_{[y_j,y_{j+1}]}(y)\,dy\right) = \\ &= C_{i,j}\left(\int_a^b \chi_{[x_i,x_{i+1}]}(x)\,dx\right)\cdot\left(\int_c^d \chi_{[y_j,y_{j+1}]}(y)\,dy\right), \end{aligned} \tag{21.578}$$

et utiliser les propriétés de linéarité de l'intégrale

$$\begin{aligned} \int_{\mathbb{R}^2} f\,dV &= \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\left(\int_a^b \chi_{[x_i,x_{i+1}]}(x)\,dx\right)\cdot\left(\int_c^d \chi_{[y_j,y_{j+1}]}(y)\,dy\right) = \\ &= \int_c^d\int_a^b \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{[x_i,x_{i+1}]}(x)\cdot\chi_{[y_j,y_{j+1}]}(y)\,dxdy = \\ &= \int_c^d\int_a^b f\,dxdy. \end{aligned} \tag{21.579}$$

De même on obtient

$$\begin{aligned} \int_{\mathbb{R}^2} f\,dV &= \int_a^b\int_c^d \sum_{j=0}^{n-1}\sum_{i=0}^{m-1} C_{i,j}\,\chi_{[x_i,x_{i+1}]}(x)\cdot\chi_{[y_j,y_{j+1}]}(y)\,dxdy = \\ &= \int_a^b\int_c^d f\,dxdy. \end{aligned} \tag{21.580}$$

En général, on prouve la proposition suivante

Proposition 21.172.
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 $= \prod_{i=1}^{p}[a_i, b_i]$. Alors

$$\begin{aligned}\int_{\mathbb{R}^p} f(x_1,\dots,x_p)\,dV &= \int_{a_p}^{b_p}\int_{a_{p-1}}^{b_{p-1}}\cdots\int_{a_1}^{b_1} f(x_1,\dots,x_p)\,dx_1\cdots dx_p = \\ &= \int_{a_{s_p}}^{b_{s_p}}\int_{a_{s_{p-1}}}^{b_{s_{p-1}}}\cdots\int_{a_{s_1}}^{b_{s_1}} f(x_1,\dots,x_p)\,dx_1\cdots dx_p,\end{aligned} \tag{21.581}$$

pour toute permutation $(s_1,\dots,s_p)$ de l'ensemble $\{1,\dots p\}$.

21.25.6 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 : λf est intégrable pour tout réel λ

$$\int_{\mathbb{R}^p}\lambda f\,dV = \lambda\int_{\mathbb{R}^p} f\,dV,$$

Monotonie Si $f \leqslant g$ alors

$$\int_{\mathbb{R}^p} f\,dV \leqslant \int_{\mathbb{R}^p} g\,dV,$$

Inégalité fondamentale

$$\left|\int_{\mathbb{R}^p} f\,dV\right| \leqslant \int_{\mathbb{R}^p}|f|\,dV.$$

Cette dernière inégalité s'obtient de la façon suivante :

$$\left|\int_{\mathbb{R}^p} f\,dV\right| = \left|\sum_{k\in K} C_k m(R_k)\right| \leqslant \sum_{k\in K}|C_k| m(R_k) = \int_{\mathbb{R}^p}|f|\,dV.$$

Inégalité de Čebičeff Si f est une application en escalier alors pour tout $a>0$ dans $\mathbb{R}$ l'ensemble $\{x\in\mathbb{R}^p\,:\,|f(x)|\geqslant a\}$ est pavable et borné, et l'inégalité suivante est satisfaite

$$m\left(\{x\in\mathbb{R}^p\,:\,|f(x)|\geqslant a\}\right) \leqslant \frac{1}{a}\int_{\mathbb{R}^p}|f|\,dV.$$

21.25.7 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 21.173.
Soit $f:\mathbb{R}^p\to\mathbb{R}$ une fonction.

— *Pour toute application en escalier intégrable f_* telle que $f_*\leqslant f$, l'intégrale de f_* est dit une* ***somme inférieure*** *de f.*

— *Pour toute application en escalier intégrable f^* telle que $f_*\geqslant f$, l'intégrale de f^* est dit une* ***somme supérieure*** *de f.*

Soient $\sum_* f$ et $\sum^* f$ 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 $\sum_* f$ et b est dans $\sum^* f$ alors $a \leqslant b$.

Définition 21.174.
La fonction f est intégrable (au sens de Riemann) si $\sum_* f$ et $\sum^* f$ 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 21.175.
Toute fonction intégrable est bornée et à support compact. En effet, si le support de la fonction n'est pas compact alors soit $\sum_* f$ soit $\sum^* f$ 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 21.176.
Toute fonction continue à support compact est intégrable.

Cette proposition n'est a priori pas étonnante, vu qu'une fonction continue sur un support compact est bornée (théorème de Weierstrass 7.99).

21.25.8 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.

Théorème 21.177 (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 21.177. En dimension plus haute, on a le même résultat

Théorème 21.178.
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 21.179
Le premier exemple à faire est celui d'une fonction en escalier intégrable et positive. Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ la fonction

$$f(x,y) = \begin{cases} 1 & \text{si } (x,y) \in R_1 =]-1,3] \times [4,5] \\ 3 & \text{si } (x,y) \in R_2 =]13,15[\times [0,2[\\ 0 & \text{dans les autres cas.} \end{cases} \tag{21.582}$$

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. △

Exemple 21.180
On veut calculer le volume du solide S, borné par le paraboloïde elliptique $x^2 + 2y^2 + z = 16$ et le plans $x = 2$, $x = 0$, $y = 2$ $y = 0$, $z = 0$. 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 21.177 et décomposer l'intégrale double en deux intégrales simples :

$$\begin{aligned} &\int_R 16 - x^2 - 2y^2 \, dV = \int_0^2 \int_0^2 f(x,y)\, dxdy = \\ &= \int_0^2 \left[(16 - 2y^2)x - \frac{x^3}{3} \right]_{x=0}^{x=2} dy = \\ &= \left[\left(32 - \frac{8}{3} \right) y - \frac{4y^3}{3} \right]_{x=0}^{x=2} = 64 - \frac{16+32}{3} = 48. \end{aligned} \tag{21.583}$$

Vérifiez, comme exercice, qu'on obtient le même résultat en intégrant d'abord par rapport à y et puis par rapport à x. △

Exemple 21.181
Dans les hypothèses du théorème 21.177 l'ordre des intégrations partielles ne change pas la valeur de l'intégrale. En fait, si les calculs sont faites 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

$$f(x,y) = \begin{cases} y \sin(x,y) & \text{si } (x,y) \in [1,2,] \times [0,\pi] \\ 0 & \text{sinon.} \end{cases} \tag{21.584}$$

Les deux section 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

$$\int_0^\pi \cos y - \cos(2y)\, dy.$$

△

21.25.9 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 (s'il existe) de sa fonction caractéristique.

On peut dire que 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 suivantes

Régions du premier type A est borné et contenu entre les graphes de deux fonctions continues de x

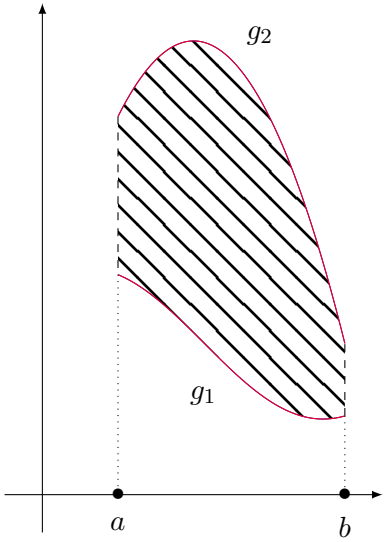
$$A = \{(x,y) \in \mathbb{R}^2 \,:\, a \leqslant x \leqslant b,\, g_1(x) \leqslant y \leqslant g_2(x)\},$$

avec g_1 et g_2 continues.

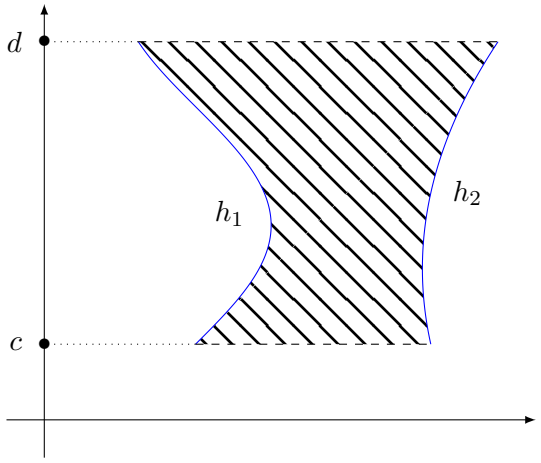
Régions du deuxième type A est borné et contenu entre les graphes de deux fonctions continues de y

$$A = \{(x,y) \in \mathbb{R}^2 \,:\, c \leqslant y \leqslant d,\, h_1(y) \leqslant x \leqslant h_2(y)\},$$

avec h_1 et h_2 continues.



(a) Une region du premier type



(b) Une region du deuxième type

FIGURE 21.19 – Régions du premier et du deuxième type

Exemple 21.182
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 admets les représentations suivantes

$$C=\{(x,y)\in\mathbb{R}^2\,:\,0\leqslant x\leqslant 1,\,x^2\leqslant y\leqslant 2x\},$$

et

$$C=\{(x,y)\in\mathbb{R}^2\,:\,0\leqslant y\leqslant 1,\,y/2\leqslant x\leqslant \sqrt{y}\}.$$

△

Définition 21.183.
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

$$\bar{f}(x,y)=\begin{cases} f(x,y), & si\ (x,y)\in A,\\ 0, & sinon.\end{cases} \tag{21.585}$$

*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

$$\begin{aligned}\int_A f\,dV&=\int_{\mathbb{R}^2}\bar{f}\,dV=\\&=\int_a^b\int_c^d\bar{f}\,dydx=\\&=\int_a^b\left(\int_c^{g_1(x)}\bar{f}\,dy+\int_{g_1(x)}^{g_2(x)}\bar{f}\,dy+\int_{g_2(x)}^{d}\bar{f}\,dy\right)dx=\\&=\int_a^b\int_{g_1(x)}^{g_2(x)}f\,dydx.\end{aligned} \tag{21.586}$$

De même, si A est du deuxième type on obtient

$$\int_A f\,dV=\int_c^d\int_{h_1(y)}^{h_2(y)}f\,dxdy. \tag{21.587}$$

Exemple 21.184

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 : $B((0,0),1)=\{(x,y)\,:\,x\in[-1,1],\,-\sqrt{1-x^2}\leqslant y\leqslant\sqrt{1-x^2}\}$.

$$I=\int_B\sqrt{1-x^2}\,dV=\int_{-1}^1\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\sqrt{1-x^2}dydx \tag{21.588}$$

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 :

$$I=\int_{-1}^1\left[y\sqrt{1-x^2}\right]_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}}dx=2\int_{-1}^1(1-x^2)dx. \tag{21.589}$$

Cela est à nouveau une intégrale simple à effectuer. Le résultat est

$$2\int_{-1}^{1}(1-x^2)dx = 2\left[x-\frac{x^3}{3}\right]_{x=-1}^{x=1} = \frac{8}{3}. \tag{21.590}$$

△

Remarque 21.185.
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.

21.186.
Par le lemme 15.157 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 22.11.

Exemple 21.187
On veut calculer l'aire de la région de la figure 21.20 définie par

$$A = \{(x,y) \in \mathbb{R}^2 \,|\, 0 \leqslant x \leqslant 1, x^3 - 1 \leqslant y \leqslant x\}.$$

On considère l'intégrale

$$\int_{\mathbb{R}^2} \chi_A \, dV = \int_0^1 \int_{x^3+1}^{x} 1 \, dy \, dx = \int_0^1 -x^3 + x + 1 \, dx = -\frac{1}{4} + \frac{1}{2} + 1 = \frac{5}{4}.$$

△

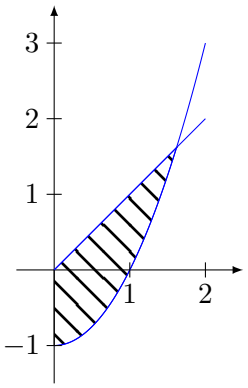


FIGURE 21.20 – La région A de l'exemple 21.187

Exemple 21.188
Parfois la région sur laquelle on veut intégrer peut être décrite indifféremment de deux façons, mais la fonction à intégrer nous force a 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 21.189.
Soit A un sous-ensemble borné de $\mathbb{R}^2$ et soient B_1 et B_2 deux parties de A telles que $B_1 \cap B_2 = \varnothing$

et $B_1 \cup B_2 = A$. Alors, pour toute fonction f intégrable sur A (et en particulier pour sa fonction caractéristique) on a

$$\int_A f\,dV = \int_{B_1} f\,dV + \int_{B_2} f\,dV.$$

Exemple 21.190
La région D que nous voyons sur la figure 21.21 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,

$$D_1 = \{(x, y) \,:\, -3 \leqslant x \leqslant -1,\, -\sqrt{2x+6} \leqslant y \leqslant \sqrt{2x+6}\},$$

et

$$D_2 = \{(x, y) \,:\, -3 \leqslant x \leqslant -1,\, x - 1 \leqslant y \leqslant \sqrt{2x+6}\}.$$

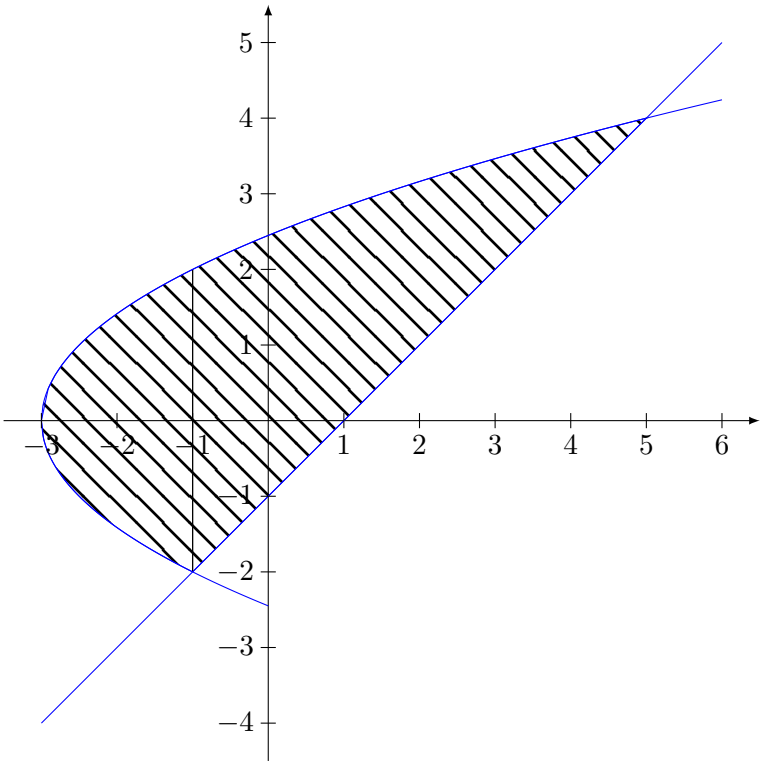


FIGURE 21.21 – La région D de l'exemple 21.190

△

21.25.10 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.

Définition 21.191.
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.*

$$E = \{(x, y, z) \in \mathbb{R}^3 \,|\, (x, y) \in A \subset \mathbb{R}^2, u_1(x, y) \leqslant z \leqslant u_2(x, y)\}.$$

Le sous-ensemble de A de $\mathbb{R}^2$ qui apparaît dans la définition 21.191 est la projection (ou l'ombre) de E sur le plan x-y.

Exemple 21.192
La région E donnée par une portion de sphère collée à un cône est une région solide de premier type

$$E = \{(x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in \overline{B\big((0,0),1\big)}, \sqrt{x^2+y^2} \leqslant z \leqslant \sqrt{1-x^2-y^2}\}. \tag{21.591}$$

L'ombre de E est la boule unité de $\mathbb{R}^2$. L'ensemble $\sqrt{x^2+y^2} \leqslant z$ est un cône posé sur sa pointe tandis que l'ensemble $z \leqslant \sqrt{1-x^2-y^2}$ est la demi-sphère. L'ensemble E contient les points entre les deux, voir la figure 21.22.

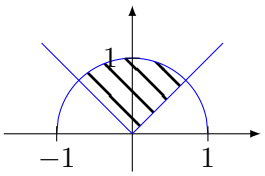


FIGURE 21.22 – Il faut voir ça en trois dimensions.

△

Si la fonction f, à intégrer sur E, et ses sections sont intégrables alors on peut réduire l'intégrale

$$\begin{aligned}\int_E f(x,y,z)\,dV &= \int_A \left(\int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\,dz \right) dV = \\ &= \int_A \left(F(x,y,u_2(x,y)) - F(x,y,u_1(x,y)) \right) dV,\end{aligned} \tag{21.592}$$

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 21.193.
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

$$\int_A \left(u_2(x,y) - u_1(x,y) \right) dV.$$

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 21.194
On veut calculer le volume de la boule de rayon a, centrée à l'origine $B = \{(x,y,z) \in \mathbb{R}^3 \,|\, x^2 + y^2 + z^2 \leqslant a^2\}$. On peut décrire B par

$$B = \left\{ (x,y,z) \in \mathbb{R}^3 \,|\, (x,y) \in D_a, -\sqrt{a^2-x^2-y^2} \leqslant z \leqslant \sqrt{a^2-x^2-y^2} \right\},$$

où D_a est le disque de rayon a centré en $(0,0)$, donc le volume B sera

$$2 \int_{D_a} \sqrt{a^2-x^2-y^2} dV.$$

Cet 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 $\sqrt{a^2-z^2}$. L'aire d'un tel disque est $\pi(a^2+z^2)$. 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 21.195**

On calcul l'intégrale de $f(x,y,z)=z$ sur la pyramide P bornée par le plans $x=0$, $y=0$, $x+y+z=1$, $x+y+z/2=1$. On remarque tout de suite que le 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

$$P=\{(x,y,z)\in\mathbb{R}^3\,|\,(x,y)\in T,\,1-2x-2y\leqslant z\leqslant 1-x-y\}$$

et T par

$$T=\{(x,y)\in\mathbb{R}^2\,|\,0\leqslant x\leqslant 1,\,0\leqslant y\leqslant 1-x\},$$

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.

21.25.11 Intégrales de fonctions non bornées sur des ensembles non bornés

Soit $f\colon\mathbb{R}^n\to\overline{\mathbb{R}}$, une fonction positive. On dit qu'elle est **intégrable** sur $E\subset\mathbb{R}^n$ si

(1) $\forall r>0$, la fonction $f_r(x)=f(x)\mathbb{1}_{f<r}$ est intégrable sur E_r ;

(2) la limite $\lim_{r\to\infty}\int_{E_r} f_r$ est finie.

Dans ce cas, on pose

$$\int_E f=\lim_{r\to\infty}\int_{E_r} f_r. \tag{21.593}$$

Théorème 21.196.
Soit E mesurable dans $\mathbb{R}^n$ et $f\colon E\to\overline{\mathbb{R}}$. Si f est mesurable et s'il existe $g\colon E\to\overline{\mathbb{R}}$ intégrable sur E telle que $|f(x)|\leqslant g(x)$ pour tout $x\in E$, alors f est intégrable sur E.

Réciproquement, si f est intégrable sur E, alors f est mesurable.

Lemme 21.197.
Si f est une fonction sur $[a,\infty[$, alors nous avons la formule

$$\lim_{b\to\infty}\int_a^b f(x)dx=\int_a^\infty f(x)dx \tag{21.594}$$

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

$$\psi(x)=\int_a^x f \tag{21.595}$$

est bornée. Soit M, un majorant. Pour toute fonction simple φ dominant f, on a $\int\varphi\leqslant M$, 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

$$\int_a^\infty f\leqslant\lim_{b\to\infty}\int_a^b f. \tag{21.596}$$

□

21.26 Intégrale de Wallis

Lemme 21.198 ([361]).
Soit $n>0$. En posant $I_n=\int\sin^n(x)dx$, nous avons :

$$I_n=\frac{\cos(x)\sin^{n-1}(x)}{n}+\frac{n-1}{n}I_{n-2}, \tag{21.597}$$

c'est à dire

$$\int\sin^n(x)dx=\frac{n-1}{n}\int\sin^{n-2}(t)dt-\frac{\sin^{n-1}(x)\cos(x)}{n}. \tag{21.598}$$

De la même manière,

$$\int\cos^n(x)dx=\frac{\cos^{n-1}(x)\sin(x)}{n}+\frac{n-1}{n}\int\cos^{n-2}(x)dx. \tag{21.599}$$

Démonstration. Nous posons $I_n=\int\sin^n(t)dt$, et nous y allons par récurrence. D'abord pour $n=1$. Dans ce cas, la formule à démontrer se réduit à

$$\int\sin(x)dx=\cos(x). \tag{21.600}$$

Pas de problèmes.

Pour $n\geqslant 2$, nous évaluons l'intégrale I_n en utilisant une intégration par partie [31] en posant

$$\begin{cases} u=\sin^{n-1}(t) & \text{(21.601a)}\\ v'=\sin(t) & \text{(21.601b)}\end{cases}$$

et en déduisant

$$\begin{cases} u'=(n-1)\sin^{n-2}(t)\cos(t) & \text{(21.602a)}\\ v=-\cos(t). & \text{(21.602b)}\end{cases}$$

Nous avons alors le calcul

$$\begin{aligned} I_n&=-\cos(x)\sin^{n-1}(x)+\int(n-1)\sin^{n-2}(x)\cos^2(x)dx && \text{(21.603a)}\\ &=-\cos(x)\sin^{n-1}(x)+(n-1)\int\sin^{n-2}\underbrace{\cos^2(x)}_{=1-\sin^2(x)}dx && \text{(21.603b)}\\ &=\cos(x)\sin^{n-1}(x)+(n-1)\underbrace{\int\sin^{n-2}(x)dx}_{=I_{n-2}}-(n-1)\underbrace{\int\sin^n(x)\,dx}_{=I_n} && \text{(21.603c)}\end{aligned}$$

Nous avons donc déjà prouvé que

$$I_n=\cos(x)\sin^{n-1}(x)+(n-1)(I_{n-2}-I_n). \tag{21.604}$$

En isolant I_n,

$$I_n=\frac{\cos(x)\sin^{n-1}(x)}{n}+\frac{n-1}{n}I_{n-2}. \tag{21.605}$$

La formule (21.599) se démontre de la même façon. □

31. Proposition 21.132.

Lemme 21.199 ([362, 363]).
Nous posons

$$W_n = \int_0^{\pi/2} \cos^n(t)dt. \tag{21.606}$$

Alors :

(1) une formule de récurrence :

$$W_n = \frac{n-1}{n} W_{n-2}, \tag{21.607}$$

(2) et une formule un peu explicite :

$$W_{2n} = \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2}. \tag{21.608}$$

Démonstration. Le nombre W_n est seulement la seconde intégrale du lemme 21.198, évaluée entre 0 et $\pi/2$. En partant donc de (21.599), nous avons

$$W_n = \int_0^{\pi/2} \cos^n(x)dx = \left[\frac{\cos^{n-1}(x)\sin(x)}{n} \right]_0^{\pi/2} + \frac{n-1}{n} \int_0^{\pi/2} \cos^{n-2}(x)dx. \tag{21.609}$$

Le terme aux bords disparaît grâce aux valeurs trigonométriques remarquables [32]. Il reste immédiatement

$$W_n = \frac{n-1}{n} W_{n-2}. \tag{21.610}$$

À partir de là, nous démontrons (21.608) par récurrence. D'abord pour $n = 0$, c'est l'égalité $W_0 = \pi/2$ qui est correcte parce que

$$\int_0^{\pi/2} \cos^0(x)dx = \int_0^{\pi/2} 1dx. \tag{21.611}$$

Pour la récurrence elle-même,

$$W_{2(n+1)} = \frac{2n+1}{2(n+1)} W_{2n} \tag{21.612a}$$

$$= \frac{2n+1}{2n+2} \frac{(2n)!}{(2^n n!)^2} \frac{\pi}{2} \tag{21.612b}$$

$$\ldots \text{pas mal de petits calculs} \ldots \tag{21.612c}$$

$$= \frac{(2n+2)!}{\big((n+1)!2^{n+1}\big)^2} \frac{\pi}{2}. \tag{21.612d}$$

Voila. □

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 21.199. Pour les termes impairs, nous n'avons rien. Dans tous les cas, nous avons la formule de récurrence

$$W_n = \frac{n-1}{n} W_{n-2} \tag{21.613}$$

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.

32. Par exemple la proposition 19.32(10).

Lemme 21.200.
La suite donnée par

$$W_n = \int_0^{\pi/2} \cos^n(x)dx \tag{21.614}$$

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

$$\cos^{n+1}(x) \leqslant \cos^n(x). \tag{21.615}$$

Les intégrales suivent les mêmes inégalités. □

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.

Lemme 21.201.
Nous posons

$$W_n = \int_0^{\pi/2} \cos^n(x)dx \tag{21.616}$$

La fonction $\alpha\colon \mathbb{N} \to \mathbb{R}$ *définie par* $W_n = \alpha(n)W_{n-1}$ *vérifie* $\lim_{n\to\infty} \alpha(n) = 1$.

Démonstration. Parce que nous en aurons besoin, nous triturons d'abord un peu la formule de récurrence (21.607). D'abord nous l'inversons un peu pour avoir

$$W_n = \frac{n+2}{n}W_{n+2}, \tag{21.617}$$

et ensuite nous écrivons

$$W_{n-1} = \frac{n+1}{n}W_{n+1}. \tag{21.618}$$

Ne vous posez pas de questions, ça va être utile. Le fait que la suite soit décroissante (lemme 21.200) nous permet d'écrite

$$W_{n+2} \leqslant W_n \leqslant W_{n-1}. \tag{21.619}$$

En y remplaçant W_n par $\alpha(n)W_{n+1}$ et W_{n-1} par (21.618),

$$W_{n+1} \leqslant \alpha(n)W_{n+1} \leqslant \frac{n+1}{n}W_{n+1}. \tag{21.620}$$

Nous simplifions par W_{n+1} et nous trouvons l'encadrement, valable pour tout n :

$$1 \leqslant \alpha(n) \leqslant \frac{n+1}{n}. \tag{21.621}$$

Nous en déduisons par la règle de l'étau que $\alpha(n) \to 1$. □

Lemme 21.202.
Soit

$$W_n = \int_0^{\pi/2} \cos^n(x)dx. \tag{21.622}$$

Encore plusieurs choses à dire.

(1) Pour tout n nous avons

$$nW_nW_{n-1} = \frac{\pi}{2}. \tag{21.623}$$

(2) Nous avons l'équivalence de suites[33]

$$W_n \sim \sqrt{\frac{\pi}{2n}}. \tag{21.624}$$

33. Définition 8.17.

Démonstration. En deux parties

La suite constante Nous posons $K_n = nW_nW_{n-1}$. Grâce aux formules de récurrence (21.607) que nous écrivons sous la forme

$$W_n = \frac{n-1}{n}W_{n-2} \tag{21.625a}$$

$$W_{n+1} = \frac{n}{n+1}W_{n-1}, \tag{21.625b}$$

nous avons

$$K_{n+1} = (n+1)W_{n+1}W_n = (n+1)\frac{n}{n+1}\frac{n-1}{n}W_{n-2}W_{n_2} = (n-1)W_{n-1}W_{n-2} = K_{n-1}. \tag{21.626}$$

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

$$W_1 = \int_0^{\pi/2} \cos(t)dt = 1. \tag{21.627}$$

Pour W_2, il ne faut pas calculer d'intégrales, mais seulement utiliser le formule (21.608). Nous trouvons vite $W_2 = \frac{\pi}{4}$. Donc

$$K_1 = W_1W_0 = \frac{\pi}{2} \tag{21.628}$$

et

$$K_2 = 2W_2W_1 = \frac{\pi}{2}. \tag{21.629}$$

L'équivalence de suites Soit la fonction α définie par $W_n = \alpha(n)W_{n-1}$. Le lemme 21.201 nous dit que $\alpha(n) \to 1$. Nous l'utilisons dans (21.623) :

$$\frac{\pi}{2} = nW_nW_{n-1} = n\alpha(n)W_{n-1}^2. \tag{21.630}$$

Donc

$$W_{n-1}^2 = \frac{\pi}{2n\alpha(n)}, \tag{21.631}$$

et

$$W_{n-1}\sqrt{\frac{1}{\alpha(n)}}\sqrt{\frac{\pi}{2n}}. \tag{21.632}$$

Vu que le coefficient $\sqrt{1/\alpha(n)}$ tend vers 1 pour $n \to 1$, nous avons l'équivalence demandée.

□

21.27 Formule de Stirling

Lemme 21.203 ([364]).
Si $n \in \mathbb{N}$ nous avons la formule

$$\ln(1+\frac{1}{n}) = 2\sum_{k=0}^{\infty}\frac{1}{2k+1}\left(\frac{1}{2n+1}\right)^{2k+1} \tag{21.633}$$

Démonstration. Il s'agit d'utiliser astucieusement le développement de la proposition 16.74. Nous avons d'une part

$$\ln(1+t) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k} \tag{21.634}$$

et d'autre part,

$$-\ln(1-t) = \sum_{k=1}^{\infty}\frac{x^k}{k}. \tag{21.635}$$

Cela permet de calculer, en utilisant l'associativité de la série [34]

$$\ln\left(\frac{1+t}{1-t}\right) = \ln(1+t) - \ln(1-t) \tag{21.636a}$$
$$= \sum_{k=1}^{\infty}\left(\frac{(-1)^{k+1}x^k}{k} + \frac{x^k}{k}\right) \tag{21.636b}$$
$$= 2\sum_{k=0}^{\infty}\frac{x^{2k+1}}{2k+1} \tag{21.636c}$$

parce que tous les termes pairs s'annulent, tandis que les termes impairs sont doublée. Notez que la somme dans (21.636c) commence à $k = 0$, contrairement aux autres qui commencent à 1.

Et là c'est l'astuce : on pose écrit l'égalité (21.636) avec le t qu'il faut pour que

$$\frac{1+t}{1-t} = 1 + \frac{1}{n}. \tag{21.637}$$

Nous posons donc $t = \frac{1}{2n+1}$ nous avons

$$\ln\left(1+\frac{1}{n}\right) = 2\sum_{k=0}^{\infty}\frac{1}{2k+1}\left(\frac{1}{2n+1}\right)^{2k+1}, \tag{21.638}$$

ce qu'il fallait. □

Lemme 21.204 (Formule de Stirling[365, 366, 362, 364, 367])**.**
Nous avons l'équivalence de suites [35]

$$n! \sim \left(\frac{n}{e}\right)^n\sqrt{2\pi n}. \tag{21.639}$$

Démonstration. Nous posons

$$a_n = \frac{n!}{\sqrt{2n}\left(\frac{n}{e}\right)^n} \tag{21.640}$$

et $b_n = \ln(a_n)$.

Une formule pour $b_n - b_{n+1}$ Nous faisons un beau calcul qui utilise les formules de la proposition 16.60 ainsi que $\ln(e) = 1$:

$$b_n - b_{n+1} = \ln\left(\frac{a_n}{a_{n+1}}\right) \tag{21.641a}$$
$$= \ln\left(\frac{(n+1)^{n+1/2}}{n^{n+1/2}}\frac{1}{e}\right) \tag{21.641b}$$
$$= \ln\left((\frac{n+1}{n})^{n+1/2}\right) - 1 \tag{21.641c}$$
$$= \left(n+\frac{1}{2}\right)\ln(1+\frac{1}{n}) - 1. \tag{21.641d}$$

34. Proposition 12.68.
35. Définition 8.17.

La suite (b_n) est décroissante Nous écrivons l'égalité (21.641) en utilisant le lemme 21.203 :

$$b_n - b_{n+1} = (n + \frac{1}{2}) \ln(1 + \frac{1}{n}) - 1 \tag{21.642a}$$
$$= \frac{1}{2}(2n+1)2 \sum_{k=0}^{\infty} \frac{1}{2k+1} \left(\frac{1}{2n+1} \right)^{2k+1} - 1 \tag{21.642b}$$
$$= (2n+1) \sum_{k=1}^{\infty} \frac{1}{2k+1} \left(\frac{1}{2n+1} \right)^{2k+1} \tag{21.642c}$$
$$= \sum_{k=1}^{\infty} \frac{1}{2k+1} \left(\frac{1}{2n+1} \right)^{2k}. \tag{21.642d}$$

Notez que le terme $k = 0$ s'est simplifié avec le -1. Vu que le tout est une somme de termes positifs, nous avons

$$b_n - b_{n+1} > 0 \tag{21.643}$$

et la suite est décroissante.

Majoration pour $b_n - b_{n+1}$ Vu que $\frac{1}{2k+1} < 1$, nous pouvons majorer :

$$b_n - b_{n+1} \leqslant \sum_{k=1}^{\infty} \left(\frac{1}{(2n+1)^2} \right)^k. \tag{21.644}$$

Nous remarquons que cela est une série géométrique déjà traitée dans l'exemple 12.75. Nous faisons un peu de calcul en partant de

$$\sum_{k=1}^{\infty} q^n = \frac{q}{1-q} \tag{21.645}$$

avec $q = \frac{1}{(2n+1)^2}$. Après quelque simplifications,

$$b_n - b_{n+1} \leqslant \frac{1}{4} \frac{1}{n(n+1)} \tag{21.646}$$

Le coup de la somme téléscopique Nous avons

$$b_1 - b_n = (b_1 - b_2) + (b_2 - b_3) + \ldots + (b_{n_1} - b_n). \tag{21.647}$$

Chacun de ces termes est majoré ; nous avons donc

$$b_1 - b_n < \frac{1}{4} \sum_{m=1}^{n-1} \frac{1}{m(m+1)} \leqslant \frac{1}{4} \sum_{m=1}^{\infty} \frac{1}{m(m+1)} = \frac{1}{4} \tag{21.648}$$

grâce au lemme 12.79 pour la dernière somme.

La suite b_n est bornée vers le bas Vu que $b_1 = 1 - \frac{\ln(2)}{2}$, nous avons

$$b_n > b_1 - \frac{1}{4} = \frac{3}{4} - \frac{\ln(2)}{2}. \tag{21.649}$$

En utilisant la majoration de l'exemple 21.148 nous trouvons

$$0.327 \leqslant b_n \leqslant 0.427. \tag{21.650}$$

Ce 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.

Une limite pour (a_n) La suite (b_n) est décroissante et bornée vers le bas, donc elle est convergente par le lemme 8.18. Vu que l'exponentielle est une fonction continue [36], la suite $a_n = e^{b_n}$ est également convergente.

Vu que $\lim_{n\to\infty} b_n > 0$, nous avons $\lim_{n\to\infty} a_n = e^{\lim_{n\to\infty} b_n} > 1$.

L'important est que nous sachions que (a_n) est une suite convergente. Nous notons L sa limite :

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2n}\left(\frac{n}{e}\right)^n} = L. \tag{21.651}$$

Ou encore :

$$\lim_{n\to\infty} \frac{n!}{L\sqrt{2n}\left(\frac{n}{e}\right)^n} = 1. \tag{21.652}$$

Autrement dit, en posant

$$n! = \alpha(n)L\sqrt{2n}\left(\frac{n}{e}\right)^n, \tag{21.653}$$

nous avons $\lim_{n\to\infty} \alpha(n) = 1$.

Introduction de Wallis Nous avons déjà parlé dans le lemme 21.199 du nombre

$$W_{2n} = \frac{(2n)!}{(2^n n!)^2}\frac{\pi}{2}. \tag{21.654}$$

Le lemme 21.202 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 (21.653). Après pas mal de calculs [37], nous trouvons

$$I_n = \frac{\alpha(2n)}{\alpha(n)^2}\frac{1}{\sqrt{n}}\frac{\pi}{2n}. \tag{21.655}$$

Nous avons donc

$$\sqrt{\frac{\pi}{4n}} \sim \frac{\alpha(2n)}{\alpha(n)^2}\frac{\pi}{\sqrt{n}}\frac{1}{2L}. \tag{21.656}$$

Il existe donc une fonction $\beta(n)$ avec $\beta(n) \to 1$ telle que

$$\frac{\sqrt{\pi}}{2}\frac{1}{\sqrt{n}} = \beta(n)\frac{\alpha(2n)}{\alpha(n)^2}\frac{\pi}{\sqrt{n}}\frac{1}{2L}. \tag{21.657}$$

En simplifiant par $1/\sqrt{n}$,

$$\frac{\sqrt{\pi}}{2} = \beta(n)\frac{\alpha(2n)}{\alpha(n)^2}\frac{\pi}{2L}. \tag{21.658}$$

Nous prenons à présent la limite $n \to \infty$ en nous rappelant que α et β donnent 1. Après simplifications, nous trouvons

$$L = \sqrt{\pi}. \tag{21.659}$$

La fin Nous introduisons la valeur $L = \sqrt{\pi}$ dans l'expression (21.653) de la factorielle :

$$n! = \alpha(n)\sqrt{2\pi n}\left(\frac{n}{e}\right)^n. \tag{21.660}$$

La dernière équation est exactement ce qui signifie l'équivalence de suite demandée. □

36. Parce qu'elle est dérivable, voir par exemple le théorème 16.54.
37. Si vous êtes en manque de papier de brouillon, c'est le moment de vous inquiéter.

Chapitre 22

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 19.32, et c'est la proposition 19.58 qui nous permet de dire que $\big(\cos(t),\sin(t)\big)$ 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.

22.1 Définitions

Définition 22.1.
*Un **arc paramétré** dans $\mathbb{R}^p$ est un couple (I,γ) où I est un intervalle de $\mathbb{R}$ et γ est une application continue de I dans $\mathbb{R}^p$. Nous disons que (I,γ) 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,γ) est un arc de la forme (I_0,γ) avec $I_0\subset I$.

Définition 22.2.
*Un **chemin** dans $\mathbb{R}$ est une application continue*

$$\begin{aligned}\sigma\colon [a,b]&\to\mathbb{R}^3\\ t&\mapsto\sigma(t).\end{aligned}\tag{22.1}$$

La fonction $\sigma'(t)$ est la **vitesse** du chemin σ. Si la fonction $t\mapsto\sigma(t)$ est dérivable, on dit que $\sigma''(t)$ est l'**accélération**. Les points $\sigma(a)$ et $\sigma(b)$ sont les extrémités du chemin. L'ensemble

$$\{\sigma(t)\text{ tel que }t\in[a,b]\}\tag{22.2}$$

est la **courbe** σ.

22.2 Longueur d'arc

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.

Définition 22.3.
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*

$$a = t_0 < t_1 < \ldots < t_n = b. \tag{22.3}$$

Nous disons qu'une subdivision σ' est ***plus fine*** *que la subdivision σ si l'ensemble des points de σ est inclus dans celui des points de σ'. Dans ce cas, la subdivision σ' est un* ***raffinement*** *de σ. 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*.

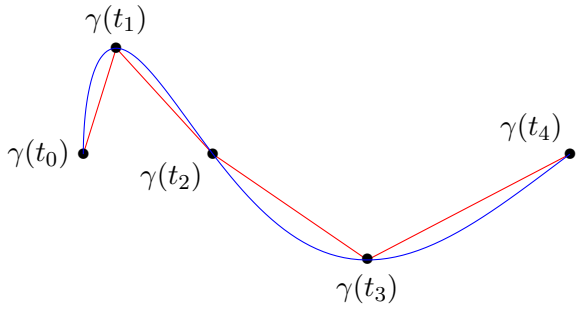


FIGURE 22.1 – La longueur d'un découpage. La somme des longueurs des segments droits est facile à calculer.

Définition 22.4.
Soit un arc paramétré compact (I, γ) et une subdivision $\sigma = (t_i)_{i=1}^n$ de $I = [a, b]$. À partir de γ et du découpage σ nous définissons le nombre (voir figure 22.1)

$$l_\sigma(\gamma) = \sum_{i=1}^{n} \|\gamma(t_i) - \gamma(t_{i-1})\|. \tag{22.4}$$

On appelle ***longueur*** *de l'arc γ le nombre*

$$l(\gamma) = \sup_\sigma l_\sigma \in [0, \infty]. \tag{22.5}$$

Nous disons que γ 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 σ_1 est plus fine que σ, nous avons

$$l_\sigma(\gamma) \leqslant l_{\sigma_1}(\gamma), \tag{22.6}$$

Comme cela peut être vu sur la figure 22.2.

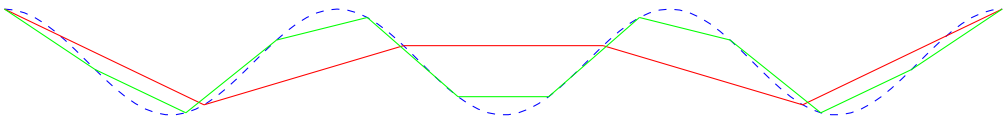


FIGURE 22.2 – Il est visible que la longueur donnée par l'approximation par des petits segments (verts) est plus longue et plus précise que celle donnée par les longs segments (rouge).

Proposition 22.5.
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 γ est un arc paramétré joignant P et Q, la longueur de γ est donné par un supremum dont un des éléments est la longueur du segment de droite. □

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 22.10).

Lemme 22.6.
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 γ 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 (22.4) 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 $l(\gamma) \geqslant l_\sigma(\gamma) > 0$. □

Proposition 22.7.
Soit (I, γ) un arc paramétré compact.

(1) Si $\gamma' = (I', \gamma)$ avec $I' \subset I$, alors $l(\gamma') \leqslant l(\gamma)$.

(2) Soit $c \in [a, b]$, et considérons les arcs $\gamma_1 = \big([a, c], \gamma\big)$ et $\gamma_2 = \big([c, b], \gamma\big)$. Alors

$$l(\gamma) = l(\gamma_1) + l(\gamma_2). \tag{22.7}$$

En particulier, γ est rectifiable si et seulement si γ_1 et γ_2 le sont.

Démonstration. (1) Nous notons $I = [a, b]$ et $I' = [a', b']$. Étant donné que $I' \subset I$, nous avons

$$a \leqslant a' < b' \leqslant b. \tag{22.8}$$

Pour chaque subdivision $\sigma_0 : a' = t_0 < t_1 < \ldots < t_n = b'$ de I', nous pouvons construire une subdivision de I en « ajoutant » les points a et b, c'est-à-dire

$$\sigma : a \leqslant t_0 < \ldots < t_n \leqslant b. \tag{22.9}$$

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

$$l_{\sigma_0}(\gamma') \leqslant l_\sigma(\gamma) \leqslant \sup_\sigma l_\sigma(\gamma) = l(\gamma). \tag{22.10}$$

Étant donné que pour toute subdivision σ_0 nous avons $l_{\sigma_0}(\gamma') \leqslant l(\gamma)$, en prenant le supremum sur les subdivisions σ_0 de I', nous avons comme annoncé

$$l(\gamma') \leqslant l(\gamma). \tag{22.11}$$

(2) Soit $\sigma = \{t_i\}$ une subdivision de $[a, b]$. Nous considérons les subdivisions σ_1 et σ_2 définies comme suit :

$$\begin{aligned} \sigma_1 &: \{t_i \text{ tel que } t_i < c\} \cup \{c\}, \\ \sigma_2 &: \{t_i \text{ tel que } t_i > c\} \cup \{c\}. \end{aligned} \tag{22.12}$$

L'inégalité triangulaire implique que

$$l_\sigma(\gamma) \leqslant l_{\sigma\cup\{c\}}(\gamma) = l_{\sigma_1}(\gamma_1) + l_{\sigma_2}(\gamma_2) \leqslant l(\gamma_1) + l(\gamma_2). \tag{22.13}$$

Nous avons donc

$$l(\gamma) \leqslant l(\gamma_1) + l(\gamma_2). \tag{22.14}$$

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 σ_1 parcours toutes les subdivisions possibles, il existe une partition σ_1^ε telle que (idem pour γ_2)

$$\begin{aligned} l_{\sigma_1^\varepsilon}(\gamma_1) + \frac{\varepsilon}{2} > l(\gamma_1), \\ l_{\sigma_2^\varepsilon}(\gamma_2) + \frac{\varepsilon}{2} > l(\gamma_2), \end{aligned} \tag{22.15}$$

où σ_1^ε est une subdivision de $[a,c]$ et σ_2^ε en est une de $[c,b]$. En faisant la somme des deux équations (22.15), nous trouvons

$$l(\gamma_1) + l(\gamma_2) < l_{\sigma_1^\varepsilon}(\gamma_1) + l_{\sigma_2^\varepsilon}(\gamma_2) + \varepsilon = l_{\sigma_1^\varepsilon \cup \sigma_2^\varepsilon}(\gamma) \leqslant l(\gamma) + \varepsilon. \tag{22.16}$$

L'inégalité $l(\gamma_1) + l(\gamma_2) < l(\gamma) + \varepsilon$ étant valable pour tout ε, nous avons

$$l(\gamma_1) + l(\gamma_2) \leqslant l(\gamma). \tag{22.17}$$

Cette inégalité, combinée avec l'inégalité (22.14), donne bien $l(\gamma) = l(\gamma_1) + l(\gamma_2)$.

□

22.3 Abscisse curviligne

Définition 22.8.
Soit (I,γ) un arc rectifiable compact avec $I = [a,b]$. L'application

$$\begin{aligned} \varphi\colon [a,b] &\to \mathbb{R}^+ \\ t &\mapsto l\big([a,t],\gamma\big) \end{aligned} \tag{22.18}$$

*est la **longueur d'arc** de γ.*

Cette fonction nous permet de calculer la distance (suivant la courbe) entre deux points arbitraires parce que si $a \leqslant t < u \leqslant b$, nous avons

$$l\big([t,u],\gamma\big) = \varphi(u) - \varphi(t). \tag{22.19}$$

En effet,

$$\varphi(u) - \varphi(t) = l\big([a,u],\gamma\big) - l\big([a,t],\gamma\big), \tag{22.20}$$

mais en utilisant la proposition 22.7, nous avons

$$l\big([a,u],\gamma\big) = l\big([a,t],\gamma\big) + l\big([t,u],\gamma\big). \tag{22.21}$$

Proposition 22.9.
La longueur d'arc d'un arc rectifiable compact est une fonction continue et croissante.

Démonstration. Soit (I,γ) un arc paramétré rectifiable compact avec $I = [a,b]$. Afin de montrer que φ est croissante, prenons $t \in I$ ainsi que $h > 0$ et montrons que $\varphi(t+h) \geqslant \varphi(t)$. La proposition 22.7 implique que

$$l\big([a,t+h],\gamma\big) = l\big([a,t],\gamma\big) + l\big([t,t+h],\gamma\big), \tag{22.22}$$

c'est-à-dire

$$\varphi(t+h) = \varphi(t) + l\big([t,t+h],\gamma\big) \geqslant \varphi(t). \tag{22.23}$$

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

$$|\varphi(t+s) - \varphi(t)| \leqslant \varepsilon, \qquad \forall t \in [a,b].$$

Étant donné que $l\big([t,b],\gamma\big)$ est le supremum des $l_\sigma\big([t,b],\gamma\big)$, il existe une subdivision σ donnée par les points $t,t_1,\cdots,t_{n-1},b$ telle que

$$l_\sigma\big([t,b],\gamma\big) > l\big([t,b],\gamma\big) - \frac{\varepsilon}{2} = \varphi(b) - \varphi(t) - \frac{\varepsilon}{2}. \tag{22.24}$$

La continuité de γ implique qu'il existe un η tel que

$$s \in [0,\eta] \Rightarrow \|\gamma(t+s) - \gamma(t)\| < \frac{\varepsilon}{2} \tag{22.25}$$

Quitte à prendre η 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 σ' est plus fine que σ, le nombre $l_\sigma\big([t,b],\gamma\big)$ est inférieur ou égal à $l_{\sigma'}\big([t,b],\gamma\big)$. Nous avons donc les inégalités

$$\begin{aligned}
\varphi(b) - \varphi(t) - \frac{\varepsilon}{2} &\leqslant l_\sigma\big([t,b],\gamma\big) \\
&\leqslant l_{\sigma'}\big([t,b],\gamma\big) \\
&= \big\|\gamma(t+s) - \gamma(t)\big\| + l_{\sigma'\setminus\{t\}}\big([t+s,b]\gamma\big) \\
&\leqslant \|\gamma(t+s) - \gamma(t)\| + \varphi(b) - \varphi(t+s) \\
&\leqslant \frac{\varepsilon}{2} + \varphi(b) - \varphi(t+s).
\end{aligned} \tag{22.26}$$

Au final, nous avons trouvé que

$$\varphi(t+s) - \varphi(t) \leqslant \varepsilon, \tag{22.27}$$

ce qui prouve que φ est continue au point t. □

En guise de paramètre sur un arc, nous pouvons utiliser la longueur d'arc elle-même. En effet si (I,γ) est un arc de longueur l, nous pouvons donner le même arc avec le couple $\big([0,l],g\big)$ où g est la fonction qui au réel s fait correspondre l'élément $\gamma\big(\varphi^{-1}(s)\big)$ de $\mathbb{R}^n$. Dire

$$P = (\gamma \circ \varphi^{-1})(s) \tag{22.28}$$

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 22.7.

22.3.1 Formule intégrale de la longueur

Nous pouvons voir un chemin γ 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 *scalaire* 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 γ :

$$l(\gamma) = \int_a^b \|\gamma'(t)\|dt. \tag{22.29}$$

Plus explicitement, si $\gamma(t) = \big(x(t),y(t),z(t)\big)$, alors nous aurions la formule

$$l(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}dt. \tag{22.30}$$

Théorème 22.10.
Soit (I,γ) un arc paramétré compact de classe $\mathcal{C}^1$. Alors γ est rectifiable et

$$l(\gamma)=\int_a^b\|\gamma'(t)\|dt=\int_\gamma 1, \tag{22.31}$$

où $I=[a,b]$.

Démonstration. L'égalité avec l'intégrale le long de γ de la fonction 1 est simplement la définition 21.40 de l'intégrale curviligne.

Si $\sigma=\{t_i\}$ est une subdivision de l'intervalle $[a,b]$, alors

$$\begin{aligned} l_\sigma(\gamma)&=\sum_{i=1}^n\|\gamma(t_i)-\gamma(t_{i-1})\|\\ &=\sum_{i=1}^n\|\int_{t_{i-1}}^{t_i}\gamma'(t)dt\|\\ &\leqslant\sum_{i=1}^n\int_{t_{i-1}}^{t_i}\|\gamma'(t)\|dt\\ &=\int_a^b\|\gamma'(t)\|dt. \end{aligned} \tag{22.32}$$

Cela prouve déjà que

$$l(\gamma)=\sup_\sigma l_\sigma(\gamma)\leqslant\int_a^b\|\gamma'(t)\|dt. \tag{22.33}$$

Nous devons maintenant prouver l'inégalité inverse.

Notons φ l'abscisse curviligne $\varphi(t)=l\big([a,t],\gamma\big)$. Cette dernière vérifie

$$\varphi(t+h)-\varphi(t)=l\big([t,t+h],\gamma\big)\geqslant\|\gamma(t+h)-\gamma(t)\|, \tag{22.34}$$

et en particulier

$$\left\|\frac{\gamma(t+h)-\gamma(t)}{h}\right\|\leqslant\frac{\varphi(t+h)-\varphi(t)}{h}. \tag{22.35}$$

D'autre part, en utilisant (22.33) sur le segment $[t,t+h]$, nous avons

$$\varphi(t+h)-\varphi(t)=l\big([t,t+h],\gamma\big)\leqslant\int_t^{t+h}\|\gamma'(u)\|du. \tag{22.36}$$

Cela nous permet de continuer l'inéquation (22.35) en

$$\left\|\frac{\gamma(t+h)-\gamma(t)}{h}\right\|\leqslant\frac{\varphi(t+h)-\varphi(t)}{h}\leqslant\frac{1}{h}\int_t^{t+h}\|\gamma'(u)\|du. \tag{22.37}$$

Prenons la limite $h\to 0$. 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 $u\mapsto\|\gamma'(u)\|$,

$$\lim_{h\to0}\frac{n(t+h)-n(t)}{h}=n'(t)=\|\gamma'(t)\|. \tag{22.38}$$

Au final,

$$\|\gamma'(t)\|\leqslant\varphi'(t)\leqslant\|\gamma'(t)\|, \tag{22.39}$$

c'est-à-dire $\varphi'(t)=\|\gamma'(t)\|$ et donc par le théorème fondamental du calcul intégral 15.233,

$$\varphi(t)-\varphi(a)=\int_a^t\|\gamma'(u)\|du. \tag{22.40}$$

Par construction de la longueur d'arc, $\varphi(a) = 0$ et en posant $t = b$ nous obtenons la relation recherchée :

$$l(\gamma) = \varphi(b) = \int_a^b \|\gamma'(u)\|du. \tag{22.41}$$

□

Remarque 22.11.
Cela est cohérent avec 21.186, mais il faut garder en tête que $l(\gamma)$ n'est pas le mesure de Lebesgue de l'image de γ dans $\mathbb{R}^2$. Cette dernière est nulle.

Exemple 22.12
Soient donc a et b deux points de $\mathbb{R}^m$, et γ la droite joignant a à b, c'est-à-dire

$$\gamma(t) = (1-t)a + tb \tag{22.42}$$

avec $t \in [0,1]$. Le théorème 22.10 nous enseigne que la longueur de ce chemin est

$$l\big([0,1],\gamma\big) = \int_0^1 \|\gamma'(t)\|dt = \int_0^1 \| - a + b\| = \|b-a\|, \tag{22.43}$$

qui est bien la distance entre a et b. △

Exemple 22.13(Circonférence du cercle)
Nous savons que l'image de

$$\begin{aligned}\gamma\colon [0,2\pi[&\to \mathbb{R}^2\\ t &\mapsto \big(R\cos(t), R\sin(t)\big)\end{aligned} \tag{22.44}$$

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^∞) par la proposition 19.16. La longueur sera, d'après la formule (22.10)

$$l_\gamma = \int_0^{2\pi} \|\gamma'(t)\|dt = 2\pi R \tag{22.45}$$

grâce à la formule $\sin^2 + \cos^2 = 1$ du lemme 19.18.

Mais tout cela n'est pas satisfaisant parce que nous n'avons pas encore de valeur numérique de π.

Il y a une autre façon de faire en considérant le quart de cercle dont la longueur en fonction de π est vite calculée par

$$l_\gamma = \int_0^{pi/2} \|\gamma'(t)\|dt = \frac{\pi R}{2}. \tag{22.46}$$

Cette même longueur est calculée en termes de fonctions plus courantes avec le chemin

$$\begin{aligned}\sigma\colon]0,1[&\to \mathbb{R}^2\\ t &\mapsto \left(\frac{t}{\sqrt{R^2-t^2}}\right)\end{aligned} \tag{22.47}$$

La longueur s'exprime avec

$$l_\sigma = \int_0^1 \sqrt{\frac{R^2}{R^2+t^2}}dt. \tag{22.48}$$

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 π, il est loisible de calculer une approximation numérique de l'intégrale (22.48) (avec $R = 1$) et de l'égaler à $\pi/2$. △

Exemple 22.14
Considérons l'arc de cercle de rayon R interceptée par l'angle θ présenté sur la figure 22.3.

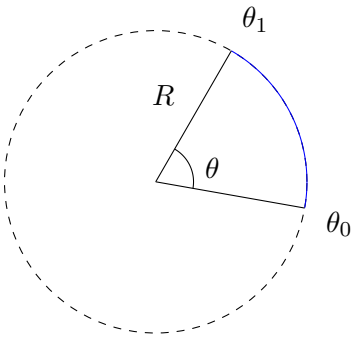


FIGURE 22.3 – Quelle est la longueur de la partie bleue de ce cercle de rayon R ?

Par définition, cette longueur sera

$$\int_{\theta_0}^{\theta_1} \sqrt{R^2\sin^2(t) + R^2\cos^2(t)}dt = R(\theta_1 - \theta_0). \tag{22.49}$$

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 [1].

Proposition 22.15.
Pour tout $x, y \in \mathbb{R}$, nous avons

$$|e^{ix} - e^{iy}| \leqslant |x - y|. \tag{22.50}$$

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 $\big(\cos(x), \sin(x)\big)$ 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

$$\begin{aligned}\sigma_1 \colon [0,1] &\to \mathbb{C}\\ t &\mapsto te^{ix} + (1-t)e^{iy}.\end{aligned} \tag{22.51}$$

Le second est l'arc de cercle

$$\begin{aligned}\sigma_1 \colon [0,1] &\to \mathbb{C}\\ t &\mapsto e^{i\left(tx+(1-t)y\right)}.\end{aligned} \tag{22.52}$$

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 22.10 :

$$l(\sigma_1) = \int_0^1 |\sigma_1'(t)| = |e^{ix} - e^{iy}|. \tag{22.53}$$

En ce qui concerne le second chemin,

$$\sigma_2'(t) = (x - y)e^{i\left(tx+(1-t)y\right)}. \tag{22.54}$$

Nous avons [2] $|\sigma_2'(t)| = |x - y|$ qui ne dépend pas non plus de t. Donc

$$l(\sigma_2) = |x - y|. \tag{22.55}$$

1. 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.
2. Si vous voulez citer des résultats, lemme 19.25 et proposition 1.131.

Étant donné la proposition 22.5 qui dit que le chemin le plus court est le segment de droite,

$$l(\sigma_1) < l(\sigma_2) \tag{22.56}$$

et donc le résultat annoncé. □

22.16.
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

$$\gamma(t) = (t, y(t)), \tag{22.57}$$

et le vecteur tangent au chemin est $\gamma'(t) = (1, y'(t))$. Donc

$$\|\gamma'(t)\| = \sqrt{1 + y'(t)^2}, \tag{22.58}$$

et

$$L = \int_a^b \sqrt{1 + y'(t)^2}. \tag{22.59}$$

Exemple 22.17
La longueur de l'hélice

$$\sigma(t) = \begin{pmatrix} \cos(2t) \\ \sin(2t) \\ \sqrt{5}t \end{pmatrix} \tag{22.60}$$

pour $t \in [0, 2\pi]$ est donnée par

$$l(\sigma) = \int_0^{4\pi} \sqrt{4\sin^2(2t) + 4\cos^2(2t) + 5}dt = \int_0^{4\pi} \sqrt{9} = 12\pi. \tag{22.61}$$

△

Définition 22.18.
*Soit $\sigma_1 \colon [a, b] \to \mathbb{R}^3$, un chemin et $\sigma_2 \colon [c, d] \to \mathbb{R}^3$, un autre chemin. On dit que ces chemins sont **équivalents** s'il existe une fonction $\varphi \colon [a, b] \to [c, d]$ strictement croissante telle que $\sigma_1(t) = \sigma_2\big(\varphi(t)\big)$.*

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 **opposée** si la fonction φ 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 paramétrisation** pour la même courbe.

Dans le cas d'une paramétrisation 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 22.19.
La longueur d'une courbe ne dépend pas du paramètre (équivalent ou opposé) choisi.

Démonstration. Soient $\sigma_1 \colon [a, b] \to \mathbb{R}^3$ et $\sigma_2 \colon [c, d] \to \mathbb{R}^3$ tels que

$$\sigma_1(t) = \sigma_2\big(\varphi(t)\big) \tag{22.62}$$

où $\varphi \colon [a, b] \to [a, d]$ est une bijection strictement monotone. Par définition on a

$$l(\sigma_1) = \int_a^b \|\sigma_1'(t)\|dt. \tag{22.63}$$

Nous pouvons exprimer la dérivée de σ_1 en termes de celle de σ_2 en dérivant la relation (22.62) :

$$\sigma_1'(t) = \varphi'(t)\sigma_2'\big(\varphi(t)\big). \tag{22.64}$$

En ce qui concerne la norme,

$$\|\sigma_1'(t)\| = |\varphi'(t)|\|\sigma_2'(t)\|. \tag{22.65}$$

Notez dans cette relation que $\varphi'(t)$ est un nombre (et non un vecteur). Étant donné que nous avons supposé que φ 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

$$\begin{aligned} l(\sigma_1) &= \int_a^b \varphi'(t)\|\sigma_2\big(\varphi(t)\big)\|dt \\ &= \int_{\varphi(a)}^{\varphi(b)} \|\sigma_2'(s)\|ds \\ &= \int_c^d \|\sigma_2'(s)\|ds \\ &= l(\sigma_2). \end{aligned} \tag{22.66}$$

Si nous considérons maintenant une paramétrisation strictement décroissante. Dans ce cas, $\varphi'(t) \leqslant 0$ et $\|\varphi'(t)\| = -\varphi'(t)$. Nous posons encore une fois $s = \varphi(t)$, $ds = \varphi'(t)ds$. 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 singe :

$$\begin{aligned} l(\sigma_1) &= \int_a^b \varphi'(t)\|\sigma_2\big(\varphi(t)\big)\|dt \\ &= -\int_{\varphi(a)}^{\varphi(b)} \|\sigma_2'(s)\|ds \\ &= -\int_d^c \|\sigma_2'(s)\|ds \\ &= \int_c^d \|\sigma_2'(s)\|ds \\ &= l(\sigma_2). \end{aligned} \tag{22.67}$$

Nous avons changé le signe en changeant l'ordre des bornes. □

22.4 Suite du chapitre

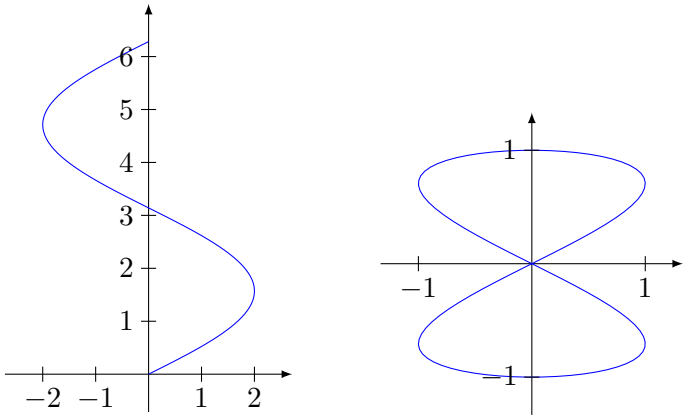
Le grand avantage des arcs paramétrés par rapports aux graphes de fonctions est le le graphe peut « faire des retours en arrière », ou bien des auto intersections. Outre les deux exemples typiques de la la figure 22.4, un exemple classique est la droite verticale. Les fonctions $y = ax+b$ permettent 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 $f : \mathbb{R} \to \mathbb{R}$ est donné par l'arc $\gamma(t) = \big(t, f(t)\big)$.

Un cercle Le cercle de rayon R est donné par l'arc $\gamma(t) = \big(R\cos(t), R\sin(t)\big)$.



(a) $\gamma(t) = (2\sin(t), t)$, avec $0 \leqslant t \leqslant 2\pi$ (b) $\gamma(t) = (sin(2t), sin(t))$ avec $-\pi \leqslant t \leqslant \pi$

FIGURE 22.4 – Des exemples d'arcs paramétrées. Ceux ne sont pas des graphes.

Remarque 22.20.
Afin d'alléger la notation, nous allons le plus souvent désigner l'arc (I, γ) simplement par la fonction γ. 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 ambigüité est à craindre.

Par exemple, lorsque nous considérons le cercle $\gamma(t) = \big(R\cos(t), R\sin(t)\big)$, 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 $I = \mathbb{R}$.

22.5 Courbes paramétrés

22.5.1 Définitions et exemples

Exemple 22.21
Soit $v \in \mathbb{R}^3$ et $x_0 \in \mathbb{R}^3$. Le chemin

$$\sigma(t) = x_0 + tv \tag{22.68}$$

est une droite. Sa vitesse est $\sigma'(t) = v$. △

Exemple 22.22
La courbe

$$\sigma(t) = \begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix} \in \mathbb{R}^2 \tag{22.69}$$

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

$$\gamma(t) = \begin{pmatrix}\cos(2\pi - t)\\ \sin(2\pi - t)\end{pmatrix} \tag{22.70}$$

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 22.23
Le chemin

$$\sigma(t) = \begin{pmatrix}t\\ t^2\end{pmatrix} \tag{22.71}$$

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 σ au point $\sigma(t)$.

Corollaire 22.24.
La tangente à un cercle est perpendiculaire au rayon.

Démonstration. Nous savons que pour un cercle,

$$y'(x) = \frac{-x}{\sqrt{R^2 - x^2}}. \tag{22.72}$$

Un point général du cercle a pour abscisse $x = R\cos(\theta)$. En remplaçant nous trouvons le coefficient directeur suivant pour la tangente :

$$y'\big(R\cos(\theta)\big) = -\frac{1}{\tan(\theta)}. \tag{22.73}$$

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 $\big(R\cos(\theta), R\sin(\theta)\big)$.

□

Exemple 22.25
Pour le cercle,

$$\sigma(t) = \begin{pmatrix}\cos(t)\\ \sin(t)\end{pmatrix}, \tag{22.74}$$

la dérivée est donnée par

$$\sigma'(t) = \begin{pmatrix}-\sin(t)\\ \cos(t).\end{pmatrix} \tag{22.75}$$

Le produit scalaire $\sigma(t) \cdot \sigma'(t)$ est nul. Le vecteur $\sigma'(t)$ est donc bien tangent (corollaire 22.24).
△

Exemple 22.26
Le courbe donnée par le chemin

$$\sigma(t) = \begin{pmatrix}\cos(t)\\ \sin(t)\\ t\end{pmatrix} \tag{22.76}$$

est une hélice. Sa vitesse est

$$\sigma'(t) = \begin{pmatrix}-\sin(t)\\ \cos(t)\\ 1\end{pmatrix}. \tag{22.77}$$

Notez que pour tout $t \in \mathbb{R}$, nous avons $\|\sigma'(t)\| = \sqrt{2}$. △

Remarque 22.27.
Lorsqu'on parle d'une courbe dans l'espace, l'intervalle sur lequel on considère la variation du paramètre est une donné fondamentale. Elle fait partie intégrante de la définition de la courbe.

22.6 Élément de longueur

22.6.1 Élément de longueur : cartésiennes

Étant donné que la longueur d'arc d'une courbe paramétrée (I, γ) 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 γ au point $\gamma(t)$.

En coordonnées cartésiennes dans le plan, une courbe paramétré est donnée par

$$\gamma(t) = \big(x_1(t), x_2(t)\big), \tag{22.78}$$

et l'élément de longueur est

$$\|x'(t)\|\,dt = \sqrt{(x_1')^2 + (x_2')^2}\,dt. \tag{22.79}$$

22.6.2 Élément de longueur : polaires (1)

En coordonnées polaires, une courbe est donnée par

$$\gamma(t) = \big(\rho(t), \theta(t)\big), \tag{22.80}$$

et le passage aux cartésiennes se fait via les formules

$$\begin{cases} x(t) = \rho(t)\cos\big(\theta(t)\big) & \text{(22.81a)}\\ y(t) = \rho(t)\sin\big(\theta(t)\big). & \text{(22.81b)} \end{cases}$$

L'élément de longueur se trouve directement en remplaçant $x(t)$ et $y(t)$ dans la formule (22.79). Les dérivées sont données par

$$\begin{aligned} x'(t) &= \rho'(t)\cos\theta(t) - \rho(t)\theta'(t)\sin\theta(t)\\ y'(t) &= \rho'(t)\sin\theta(t) + \rho(t)\theta'(t)\cos\theta(t), \end{aligned} \tag{22.82}$$

et un calcul montre que

$$\big(x'(t)\big)^2 + \big(y'(t)\big)^2 = \big(\rho'(t)\big)^2 + \big(\rho(t)\big)^2\big(\theta'(t)\big)^2. \tag{22.83}$$

Nous reviendrons plus en détail sur le concept de changement de paramétrisation (ici, les polaires) à la section 22.7.

22.6.3 Élément de longueur : polaires (2)

Parfois on utilise θ comme paramètre. L'équation de la courbe est alors donnée en coordonnées polaires sous la forme

$$\rho(\theta) = f(\theta), \tag{22.84}$$

où f est une fonction réelle et il faut comprendre que nous parlons de la courbe $\big(\rho(\theta), \theta\big)$ en coordonnées polaires. En coordonnées cartésiennes, cette courbe est donnée par

$$\begin{cases} x(t) = \rho(t)\cos(t) & \text{(22.85a)}\\ y(t) = \rho(t)\sin(t) & \text{(22.85b)} \end{cases}$$

avec t qui parcours 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. Quoi qu'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 (22.79) en y mettant les valeurs (22.85) :

$$\begin{cases} x'(t) = \rho'(t)\cos(t) - \rho(t)\sin(t) & \text{(22.86a)}\\ y'(t) = \rho'(t)\sin(t) + \rho(t)\cos(t), & \text{(22.86b)} \end{cases}$$

et

$$\big(x'(t)\big)^2 + \big(y'(t)\big)^2 = \rho'(t)^2 + \rho(t)^2. \tag{22.87}$$

Remarque 22.28.
N'oubliez pas, en utilisant ces formules, que ce qui rentre dans l'intégrale est la *racine carré* de $(x')^2 + (y')^2$.

Exemple 22.29
Calculons la circonférence du cercle. En coordonnées polaires, le graphe du cercle correspond à l'équation

$$\big(\rho(t), \theta(t)\big) = (R, t) \tag{22.88}$$

où R est constante (le rayon du cercle) et t va de 0 à 2π. En substituant dans l'équation (22.83), l'élément de longueur à intégrer est seulement

$$\sqrt{R^2} = R \tag{22.89}$$

parce que $\rho'(t) = 0$ et $\theta'(t) = 1$. La longueur du cercle est alors directement donnée par

$$l = \int_0^{2\pi} R dt = 2\pi R. \tag{22.90}$$

Nous pouvions aussi faire le calcul en coordonnées cartésiennes. Alors la courbe est donnée par les équations

$$\begin{aligned} x(t) &= R\cos(t) \\ y(t) &= R\sin(t) \end{aligned} \tag{22.91}$$

et $t \in [0, 2\pi]$. La circonférence du cercle est alors

$$l = \int_0^{2\pi} \sqrt{R^2\sin^2(t) + R^2\cos^2(t)}\, dt = \int_0^{2\pi} R\, dt = 2\pi R. \tag{22.92}$$

△

Remarque 22.30.
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. La 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 ($y > 0$).

Exemple 22.31
Une **cycloïde** est une courbe paramétrée par

$$\begin{cases} x(t) = a(t - \sin(t)) & \text{(22.93a)} \\ y(t) = a(1 - \cos(t)) & \text{(22.93b)} \end{cases}$$

avec $a > 0$ et $t \in \mathbb{R}$. Comme montré sur la figure 22.5, 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.

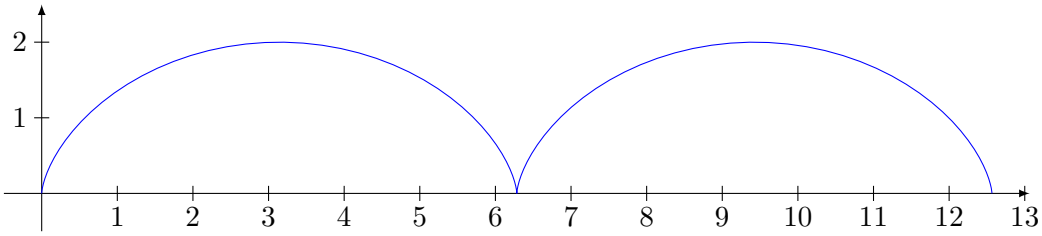


FIGURE 22.5 – La cycloïde de paramètre $a = 1$ entre 0 et 4π.

Nous avons $x'(t) = a(1 - \cos(t))$ et $y'(t) = a\sin(t)$, de telle façon que

$$\sqrt{(x')^2 + (y')^2} = a\sqrt{2 - 2\cos(t)} = a\sqrt{4\sin^2\left(\frac{t}{2}\right)} = 2a\Big|\sin\frac{t}{2}\Big|. \tag{22.94}$$

La longueur est donc donnée par

$$\int_0^{2\pi} 2a|\sin\frac{t}{2}|dt = 4a\int_0^{\pi}\sin(t)dt = 8a. \tag{22.95}$$

△

Exemple 22.32
La **cardioïde** est la courbe donnée par

$$\rho(\theta) = a(1 + \cos(\theta)). \tag{22.96}$$

avec $\theta \in [-\pi, \pi]$. Le nom de cette courbe provient de son graphe illustré à la figure 22.6.

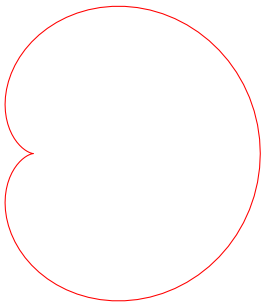


FIGURE 22.6 – Une cardioïde, $\rho = 1 + \cos(\theta)$.

L'équation (22.96) est donnée sous la forme (22.84), c'est-à-dire que $\theta(t) = t$ et $\theta'(t) = 1$, et par conséquent l'élément de longueur est donné par

$$\begin{aligned}(\rho')^2 + (\rho)^2 &= \big(-a\sin(\theta)\big)^2 + a^2\big(1 + \cos(\theta)\big)^2\\ &= a^2\sin^2(\theta) + a^2\big(1 + 2\cos(\theta) + \cos^2(\theta)\big)\\ &= a^2\big(1 + 1 + 2\cos(\theta)\big)\\ &= 2a^2\big(1 + \cos(\theta)\big)\\ &= 4a^2\cos^2\frac{\theta}{2}.\end{aligned} \tag{22.97}$$

La longueur d'arc est donc donnée par

$$l = \int_{-\pi}^{\pi} 2a\cos\frac{\theta}{2}d\theta = 2a\int_{-\pi/2}^{\pi/2}\cos(t)2dt = 8a. \tag{22.98}$$

△

22.6.4 Approximation de la longueur par des cordes

Définition 22.33.
Soit un arc paramétré (I, γ). 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 22.68 et ses dépendances, nous utilisons effectivement que l'arc γ est de classe C^2.

Nous savons que la longueur d'une courbe est donné 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 22.34.
Soit (I,γ) un arc de classe $\mathcal{C}^1$ 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

$$\left| \int_t^{t'} \|\gamma'(u)\| du - \|\gamma(t) - \gamma(t')\| \right| \leqslant 2\varepsilon |t - t'|. \tag{22.99}$$

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 γ' (parce que γ est $\mathcal{C}^1$), pour tout ε, il existe un δ tel que

$$|t - t_0| < \delta \Rightarrow \|\gamma'(t) - \gamma'(t_0)\| \leqslant \varepsilon. \tag{22.100}$$

Nous considérons la fonction

$$u \mapsto \gamma(u) - \gamma(t_0) - (u - t_0)\gamma'(t_0), \tag{22.101}$$

dont la dérivée (par rapport à u) est

$$\gamma'(u) - \gamma'(t_0). \tag{22.102}$$

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

$$\begin{aligned} \|\gamma(t) - \gamma(t_0) - (t - t_0)\gamma'(t_0) - \gamma(t') + \gamma(t_0) + (t' - t_0)\gamma'(t_0)\| & \\ = |t - t'| \|\gamma'(u) - \gamma'(t_0)\| & \\ \leqslant \varepsilon |t - t'|. & \end{aligned} \tag{22.103}$$

En simplifiant ce qui peut être simplifié dans le membre de gauche, nous trouvons

$$\|\gamma(t) - \gamma(t') - (t - t')\gamma'(t_0)\| \leqslant \varepsilon |t - t'|. \tag{22.104}$$

Le membre de gauche peut être minoré en utilisant la proposition 7.107 :

$$\Big| \|\gamma(t) - \gamma(t')\| - \|(t - t')\gamma'(t_0)\| \Big| \leqslant \varepsilon |t - t'|. \tag{22.105}$$

D'autre part, les inégalités (7.71) montrent que

$$-\|\gamma'(u) - \gamma'(t_0)\| \leqslant \|\gamma'(u)\| - \|\gamma'(t_0)\| \leqslant \|\gamma'(u) - \gamma'(t_0)\|. \tag{22.106}$$

Si de plus u est compris entre t et t', ces inégalités sont encore coincées entre $-\varepsilon$ et ε. En intégrant (22.106) par rapport à u entre t et t', nous obtenons

$$\left| \int_t^{t'} \|\gamma'(u)\| - (t - t')\|\gamma'(t_0)\| \right| \leqslant \varepsilon |t - t'|. \tag{22.107}$$

Afin d'alléger les notations pour la ligne suivante, nous notons A le nombre positif $\int_t^{t'} \|\gamma'(u)\| du$. Nous avons

$$\begin{aligned} \Big| A - \|\gamma(t) - \gamma(t')\| \Big| &= \Big| A - |t - t'| \, \|\gamma'(t_0)\| + |t - t'| \, \|\gamma'(t_0)\| - \|\gamma(t) - \gamma(t')\| \Big| \\ &\leqslant \Big| A - |t - t'| \, \|\gamma'(t_0)\| \Big| + \Big| |t - t'| \, \|\gamma'(t_0)\| - \|\gamma(t) - \gamma(t')\| \Big|. \end{aligned} \tag{22.108}$$

L'équation (22.105) 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,

$$\Big|A - |t-t'|\,\|\gamma'(t_0)\|\Big| \leqslant \Big|A - (t-t')\,\|\gamma'(t_0)\|\Big| \leqslant \varepsilon|t-t'|. \tag{22.109}$$

Au final, l'inéquation (22.108) donne

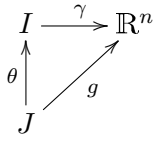
$$\Big|A - \|\gamma(t) - \gamma(t')\|\Big| \leqslant 2\varepsilon\,|t-t'|, \tag{22.110}$$

ce qu'il fallait démontrer. □

22.7 Arc géométrique

Définition 22.35.
Soient (I,γ) et (J,g) deux arcs paramétrés de classe $\mathcal{C}^k$. On dit qu'il sont ***équivalents*** *s'il existe une bijection $\theta\colon I \to J$ de classe $\mathcal{C}^k$, d'inverse de classe $\mathcal{C}^k$ telle que $g = \gamma\circ\theta$. Nous notons $\gamma \sim g$ lorsque γ et g sont équivalents (les ensembles I et J sont sous-entendus).*

Le passage d'une paramétrisation (I,γ) à une autre (J,g) se fait selon le diagramme suivant :



(22.111)

Proposition 22.36.
La relation donnée dans la définition 22.35 est une relation d'équivalence.

Démonstration. Les trois points d'une relation d'équivalence se vérifient en utilisant le fait que θ est inversible, et que l'inverse θ^{-1} jouit des mêmes propriétés de continuité ($\mathcal{C}^k$) que θ.

Réflexivité Nous avons $\gamma \sim \gamma$ avec $\theta = \mathrm{Id}$.

Symétrie Si $\gamma \sim g$, alors nous avons une application θ 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,γ) 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 22.37.
Pour cette raison les classes d'équivalences sont appelées des ***arcs géométriques*** *(de classe $\mathcal{C}^k$).*

Si Γ est une arc géométrique, ses représentants sont dits des **paramétrages admissibles** ou, plus simplement *paramétrisation*. On dit que l'application $\theta\colon J \to I$ est un **changement de variable**. Nous disons que un arc géométrique est *compact* quand ses représentants sont compacts.

Lemme 22.38.
Dans le cas d'un arc $\mathcal{C}^1$, les changements de variables sont strictement monotones (croissants ou décroissants).

Démonstration. Nous considérons (I,γ) et (J,g), deux paramétrisations différentes 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, $\theta\big(\theta^{-1}(t)\big) = t$, et par conséquent,

$$\theta'\big(\theta^{-1}(t)\big)(\theta^{-1})'(t) = 1. \tag{22.112}$$

En particulier $\theta'\big(\theta^{-1}(t)\big)$ 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 θ est strictement monotone. En effet, θ' étant continue, elle ne peut pas changer de signe sans passer par zéro (théorème 13.50 des valeurs intermédiaires). □

Théorème 22.39.
La longueur d'un arc est indépendante de sa paramétrisation, 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 22.38. Nous savons déjà que le changement de variable $\theta\colon J \to I$ est strictement monotone. Supposons que θ soit croissante. En effectuant un changement de variable dans l'intégrale qui donne la longueur [3] nous avons

$$\begin{aligned} l(\gamma) &= \int_I \|\gamma'(t)\|dt \\ &= \int_J \|\gamma'\big(\theta(s)\big)\|\theta'(s)ds \\ &= \int_J \|\gamma'\big(\theta(s)\big)\theta'(s)\|ds \\ &= \int_J \|\frac{d}{ds}(\gamma\circ\theta)(s)\|ds \\ &= \int_J \|g'(s)\|ds \\ &= l(J,g). \end{aligned} \tag{22.113}$$

□

Définition 22.40.
*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 $< \infty$.*

22.7.1 Abscisse curviligne et paramétrisation normale

Définition 22.41.
*Soit (I,γ) un arc paramétré continu rectifiable. Nous appelons **abscisse curviligne** de γ toute application $\phi\colon I \to \mathbb{R}$ telle que pour tout $t,t' \in I$ avec $t < t'$, nous ayons*

$$l\big([t,t'],\gamma\big) = \big|\phi(t') - \phi(t)\big|. \tag{22.114}$$

*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 ϕ.*

Définition 22.42.
*Un arc paramétré (I,γ_N) continu rectifiable est dit **normal** si l'identité est une abscisse curviligne.*

Lemme 22.43.
Si γ est de classe C^1 et est une paramétrisation normale, alors

(1) pour tout choix de t et t' dans I avec $t < t'$, nous avons

$$l\big([t,t'],\gamma_N\big) = t' - t. \tag{22.115}$$

(2) $\|\gamma'(t)\| = 1$ pour tout t.

Démonstration. Pour tout x_1, x_2 dans le domaine nous avons

$$l\big([x_1,x_2],\gamma\big) = \int_{x_1}^{x_2} \|\gamma'(t)\|dt = x_2 - x_1. \tag{22.116}$$

Cela implique $\|\gamma'(t)\| = 1$ pour tout t. □

3. Théorème 22.10.

Exemple 22.44
Le cercle unitaire est donné par l'arc

$$\gamma(t) = \big(\cos(t), \sin(t)\big) \tag{22.117}$$

et $t \in [0, 2\pi]$. Pour tout choix de t et t' dans $[0, 2\pi]$, nous avons

$$l\big([t, t'], \gamma\big) = \int_t^{t'} \sqrt{\sin^2(u) + \cos^2(u)}du = t' - t. \tag{22.118}$$

Les angles exprimés en radians forment donc une paramétrisation normale du cercle de rayon 1.
△

Lemme 22.45.
Pour un arc paramétré compact, la longueur d'arc est une abscisse curviligne.

Démonstration. Par définition de la longueur d'arc φ, nous avons

$$\varphi(t') - \varphi(t) = l\big([a, t'], \gamma\big) - l\big([a, t], \gamma\big) = \diamondsuit. \tag{22.119}$$

Supposons pour fixer les idées que $t' > t$. En utilisant la proposition 22.7, nous avons

$$l\big([a, t'], \gamma\big) = l\big([a, t], \gamma\big) + l\big([t, t'], \gamma\big), \tag{22.120}$$

et donc après simplification de deux termes,

$$\diamondsuit = l\big([t, t'], \gamma\big), \tag{22.121}$$

ce qui est précisément la propriété demandée pour être une abscisse curviligne. □

Proposition 22.46.
Pour tout arc paramétré C^1 sans points critiques, il existe un changement de coordonnées qui rend l'arc normal.

Démonstration. Soit (I, γ) un arc de classe $\mathcal{C}^1$. Nous devons montrer qu'il existe un intervalle J et une application $\theta\colon J \to I$ de classe $\mathcal{C}^1$ et d'inverse $\mathcal{C}^1$ tel que l'arc (J, γ_N) soit $\mathcal{C}^1$ où $\gamma_N = \gamma \circ \theta$.

Si $I = [a, b]$, nous considérons la fonction

$$\begin{aligned} \phi\colon I &\to \mathbb{R}^+ \\ t &\mapsto \int_a^t \|\gamma'(u)\|du. \end{aligned} \tag{22.122}$$

Étant définie par l'intégrale d'une fonction $\mathcal{C}^0$, la fonction ϕ est $\mathcal{C}^1$, et nous avons $\phi'(t) = \|\gamma'(t)\| > 0$ pour tout $t \in I$. Vue comme application $\phi\colon [a, b] \to [0, l(\gamma)]$, l'application ϕ est bijective et d'inverse $\mathcal{C}^1$. Voyons cela point par point.

(1) La fonction ϕ est injective parce que 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 $l \in [0, l(\gamma)]$ et $\varepsilon > 0$. Pour prouver la continuité de ϕ^{-1} en s, nous devons trouver un δ tel que

$$|s - s'| < \delta \Rightarrow \big|\phi^{-1}(s) - \phi^{-1}(s')\big| < \varepsilon. \tag{22.123}$$

Étant donné que s et s' sont dans l'image de ϕ, nous considérons les uniques t et t' tels que $s = \phi(t)$ et $s' = \phi(t')$. La quantité $\phi(t) - \phi(t')$ devient

$$\int_a^t \|\gamma'(u)\|du - \int_a^{t'} \|\gamma'(u)\|du = \int_t^{t'} \|\gamma'(u)\|du. \tag{22.124}$$

D'autre part, $\phi^{-1}(s) = t$ et $\phi^{-1}(s') = t'$, donc la condition (22.124) devient

$$|\int_{t'}^{t} \|\gamma'(u)\| du| \leqslant \delta \Rightarrow |t - t'| < \varepsilon. \tag{22.125}$$

Cela revient à la continuité des fonctions définies par une intégrale.

(4) La dérivée de son inverse est donnée par [4]

$$(\phi^{-1})'(s) = \frac{1}{\phi'\big(\phi^{-1}(s)\big)}. \tag{22.126}$$

Nous avons vu que ϕ^{-1} et ϕ' étaient continues. La fonction $(\phi^{-1})'$ étant exprimée en termes de ces deux fonctions elle est également continue.

Nous considérons l'arc paramétré (J, γ_N) avec $J = [0, l(\gamma)]$ et

$$\gamma_N(s) = (\gamma \circ \phi^{-1})(s). \tag{22.127}$$

Nous montrons maintenant que cette nouvelle paramétrisation est normale. Soient $0 \leqslant s \leqslant s' \leqslant l(\gamma)$,

$$\begin{aligned}
l\big([s, s'], g\big) &= \int_s^{s'} \|\gamma'_N(u)\| du \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|(\gamma'_N \circ \phi)(t)\| \phi'(t) dt \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|(\gamma_N \circ \phi)'(t)\| dt \\
&= \int_{\phi^{-1}(s)}^{\phi^{-1}(s')} \|\gamma'(t)\| dt \\
&= \int_0^{\phi^{-1}(s')} \|\gamma'(t)\| \, dt - \int_0^{\phi^{-1}(s)} \|\gamma'(t)\| \, dt \\
&= \phi\big(\phi^{-1}(s')\big) - \phi\big(\phi^{-1}(s)\big) \\
&= s' - s,
\end{aligned} \tag{22.128}$$

ce qui prouve que la paramétrisation (J, γ_N) est normale. □

Nous retenons que la paramétrisation normale de γ est donnée par (J, γ_N) avec $J = [0, l(\gamma)]$ et

$$\gamma_N(s) = (\gamma \circ \phi^{-1})(s) \tag{22.129}$$

où

$$\begin{aligned}
\phi\colon I &\to \mathbb{R}^+ \\
t &\mapsto \int_a^t \|\gamma'(u)\| du.
\end{aligned} \tag{22.130}$$

Notons aussi que ϕ est une fonction croissante, étant l'intégrale d'une fonction positive.

Exemple 22.47
Trouvons les coordonnées normales pour la cycloïde donnée par

$$\begin{cases} x(t) = a(t - \sin(t)), & \text{(22.131a)} \\ y(t) = a(1 - \cos(t)) & \text{(22.131b)} \end{cases}$$

et $t \in]0, 2\pi[$. Relire l'exemple 22.31.

4. Pour obtenir cette formule, dérivez les deux membres de l'équation $\phi\big(\phi^{-1}(s)\big) = s$.

D'abord nous trouvons ϕ avec la formule (22.130) avec $a = 0$. En utilisant le bout de calcul (22.94), nous avons

$$\phi(t) = 2a \int_0^t \sin \frac{u}{2} du = 4a \left(1 - \cos \frac{t}{2}\right). \tag{22.132}$$

Pour trouver $\phi^{-1}(s)$, nous résolvons l'équation

$$s = \phi\big(\phi^{-1}(s)\big) \tag{22.133}$$

par rapport à $\phi^{-1}(s)$. Dans un premier temps, nous trouvons

$$1 - \frac{s}{4a} = \cos \frac{\phi^{-1}(s)}{2}, \tag{22.134}$$

donc $\frac{\phi^{-1}(s)}{2} = \arccos(\frac{4a-s}{4a})$, et finalement

$$\phi^{-1}(s) = 2 \arccos \left(\frac{4a - s}{4a}\right). \tag{22.135}$$

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,

$$(\gamma_N)_x(s) = a\big[\phi^{-1}(s) - \sin\big(\phi^{-1}(s)\big)\big]. \tag{22.136}$$

Nous utilisons maintenant la formule trigonométrique $\sin(x) = 2 \sin \frac{x}{2} \cos \frac{x}{2}$ afin de simplifier les expressions :

$$\begin{aligned}
(\gamma_N)_x &= a\Big[2 \arccos \left(\frac{4a - s}{4a}\right) - 2 \sin \Big(\arccos \left(\frac{4a - s}{4a}\right) \Big) \cos \Big(\arccos \left(\frac{4a - s}{4a}\right) \Big)\Big] \\
&= a\Big[2 \arccos \left(\frac{4a - s}{4a}\right) - \frac{4a - s}{2a} \sqrt{1 - \left(\frac{4a - s}{4a}\right)^2}\Big] \\
&= 2a \arccos \left(\frac{4a - s}{4a}\right) - \sqrt{8as - s^2}\, \frac{4a - s}{8a}
\end{aligned} \tag{22.137}$$

où nous avons utilisé la formule $\sin\big(\arccos(x)\big) = \sqrt{1 - x^2}$. Ensuite, pour obtenir $(\gamma_N)_y$ nous devons calculer

$$(\gamma_N)_y(s) = a\big[1 - \cos\big(\phi^{-1}(s)\big)\big]. \tag{22.138}$$

Encore une fois, il est intéressant d'exprimer le cosinus en termes des angles divisés par deux : $\cos(x) = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2}$.

$$\begin{aligned}
(\gamma_N)_y &= a\Big[1 - \cos^2 \frac{\phi^{-1}(s)}{2} + \sin^2 \frac{\phi^{-1}(s)}{2}\Big] \\
&= a\Big[2 - 2 \cos^2 \frac{\phi^{-1}(s)}{2}\Big] \\
&= 2a\Big[1 - \left(\frac{4a - s}{4a}\right)^2\Big].
\end{aligned} \tag{22.139}$$

Dans cette paramétrisation, $s \in]0, 8a[$. △

Exemple 22.48

La cardioïde $\rho(\theta) = a\big(1 + \cos(\theta)\big)$ avec θ entre $-\pi$ et π. Avant d'utiliser la formule (22.130), 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 [5] (22.87) :

$$\begin{aligned}
\|\gamma'(u)\|^2 &= a^2 \sin^2(u) + a^2(1 + \cos(u))^2 \\
&= 2a^2\big(1 + \cos(u)\big),
\end{aligned} \tag{22.140}$$

5. Nous vous déconseillons d'étudier cette formule par cœur. Sachez cependant la retrouver assez vite.

et par conséquent [6]

$$\begin{aligned}\phi(t) &= \int_0^t \sqrt{2a^2\big(1+\cos(u)\big)}du\\ &= \int_0^t \sqrt{2a^2\left(1+\cos^2\frac{u}{2}-\sin^2\frac{u}{2}\right)}du\\ &= 2a\int_0^t \cos\frac{u}{2}du\\ &= 4a\sin\frac{t}{2}.\end{aligned} \tag{22.141}$$

Pour trouver l'inverse, nous résolvons $\phi\big(\phi^{-1}(s)\big)=s$ par rapport à $\phi^{-1}(s)$:

$$\begin{aligned}4a\sin\left(\frac{\phi^{-1}(s)}{2}\right)&=s,\\ \phi^{-1}(s)&=2\arcsin\left(\frac{s}{4a}\right).\end{aligned} \tag{22.142}$$

Avant d'écrire trop brutalement $\gamma_N(s)=(\gamma\circ\phi^{-1})(s)$, il faut comprendre comment est γ. Nous avons reçu la courbe sous forme polaire, c'est-à-dire

$$\gamma(t)=\big(\gamma_r(t),\gamma_\theta(t)\big)=\Big(a\big(1+\cos(t)\big),t\Big). \tag{22.143}$$

C'est comme cela qu'il faut comprendre la donnée $\rho(\theta)=a\big(1+\cos(\theta)\big)$. Maintenant la formule $\gamma_N(s)=(\gamma\circ\phi^{-1})(s)$ devient

$$\begin{cases}(\gamma_N)_r(s)=\gamma_r\big(\phi^{-1}(s)\big) & \text{(22.144a)}\\ (\gamma_N)_\theta(s)=\gamma_\theta\big(\phi^{-1}(s)\big). & \text{(22.144b)}\end{cases}$$

Étant donné que $\gamma_\theta(t)=t$, la seconde est facile :

$$(\gamma_N)_\theta(s)=2\arcsin\left(\frac{s}{4a}\right). \tag{22.145}$$

Pour la première,

$$(\gamma_N)_r(s)=a\big[1+\cos\big(2\arcsin\frac{s}{4a}\big)\big]=\frac{16a^2-s^2}{8a}. \tag{22.146}$$

Nous écrivons donc le nouveau paramétrage en coordonnées polaires sous la forme

$$\left(\frac{16a^2-s^2}{8a},2\arcsin\frac{s}{4a}\right). \tag{22.147}$$

La question qui arrive maintenant est de savoir quel intervalle parcours la nouvelle variable s. D'après le résultat de l'exemple 22.96, 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 :

$$s\in[-4a,4a], \tag{22.148}$$

ce qui est en accord avec la conditions d'existence. △

Quel enseignement tirer de cet exemple ? Lorsqu'on calcule $\phi(t)$ pour trouver les coordonnées normales, il y a deux solutions.

6. L'utilisation stricte de la formule (22.130) 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.

(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 la paramétrisation de γ. 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.

22.7.2 Tangente à une courbe paramétrée

Définition 22.49.
*Soit (I, γ) un arc paramétré de classe $\mathcal{C}^k$ avec $k \geqslant 1$. 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 $t \to t_0$.

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

$$\lim_{t\to t_0} \frac{\gamma(t) - \gamma(t_0)}{\|\gamma(t) - \gamma(t_0)\|}. \tag{22.149}$$

Évidemment si nous avions 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 $u \sim v$ s'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 22.50.
Soit (I, γ), un arc paramétré de classe $\mathcal{C}^k$ ($k \geqslant 1$) et $t_0 \in I$ tel que

$$\gamma'(t_0) = \gamma''(t_0) = \ldots = \gamma^{(q-1)}(t_0) = 0 \tag{22.150}$$

et

$$\gamma^{(q)}(t_0) \neq 0 \tag{22.151}$$

pour un entier $1 \leqslant q \leqslant k$. Alors γ 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

$$\begin{aligned}\gamma(t) = \gamma(t_0) + \gamma'(t_0)|t - t_0| + \frac{\gamma' t(t_0)}{2}|t - t_0|^2 + \cdots + \frac{\gamma^{(q)}(t_0)}{q!}|t - t_0|^q \\ + \varepsilon(t)|t - t_0|^q\end{aligned} \tag{22.152}$$

où ε est une application $\varepsilon\colon \mathbb{R} \to \mathbb{R}^n$ telle que $\lim_{t\to t_0} \varepsilon(t) = 0$. En utilisant les hypothèses, nous éliminons la majorité des termes dans le développement (22.152) :

$$\gamma(t) - \gamma(t_0) = \frac{1}{q!}\gamma^{(q)}(t_0)|t - t_0|^q + \varepsilon(t)|t - t_0|^q. \tag{22.153}$$

La direction de la droite qui joint $\gamma(t)$ à $\gamma(t_0)$ est donc donnée par

$$\frac{\gamma(t)-\gamma(t_0)}{\|\gamma(t)-\gamma(t_0)\|}=\frac{\frac{1}{q!}\gamma^{(q)}(t_0)|t-t_0|^q+\varepsilon(t)|t-t_0|^q}{\|\frac{1}{q!}\gamma^{(q)}(t_0)|t-t_0|^q+\varepsilon(t)|t-t_0|^q\|} \tag{22.154}$$

et la limite lorsque $t\to t_0$ donne $\gamma^{(q)}(t_0)$ comme direction de la tangente.

□

Lorsque le théorème s'applique, le vecteur

$$\tau=\frac{\gamma^{(q)}(t_0)}{\|\gamma^{(q)}(t_0)\|} \tag{22.155}$$

est appelé le **vecteur unitaire tangent** en $\gamma(t_0)$ à l'arc paramétré γ.

Corollaire 22.51.
Si (I,γ) 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

$$\tau(t)=\frac{\gamma'(t)}{\|\gamma'(t)\|}, \tag{22.156}$$

pour tout t dans I.

Corollaire 22.52.
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'une paramétrisation normale nous avons $\|\gamma_N'(s)\|=1$ pour tout s. Par définition,

$$l\big([s,s']\big,g\big)=\int_s^{s'}\|\gamma_N'(u)\|du=s'-s. \tag{22.157}$$

Par conséquent,

$$\lim_{h\to 0}\frac{1}{h}\int_s^{s+h}\|\gamma_N'(u)\|du=\lim_{y\to 0}\frac{s+h-s}{h}=1. \tag{22.158}$$

Cela implique que $\|\gamma_N'(s)\|=1$, et donc en particulier que (J,γ_N) est un arc régulier. Le corollaire précédent montre alors que $\tau(s)=\gamma_N'(s)/\|\gamma_N'(s)\|=\gamma_N'(s)$. □

Exemple 22.53
Considérons la courbe $\gamma(t)=(t^2,t^3)$, et cherchons la tangente en $t_0=0$. En dérivant nous avons successivement

$$\begin{aligned}\gamma(t)&=(t^2,t^3)\\ \gamma'(t)&=(2t,3t^2)\\ \gamma''(t)&=(2,6t).\end{aligned} \tag{22.159}$$

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 :

$$\frac{\gamma(t)-\gamma(t_0)}{\|\gamma(t)-\gamma(t_0)\|}=\frac{(t^2,t^3)}{\sqrt{t^4+t^6}}=\frac{(t^2,t^3)}{t^2\sqrt{1+t^2}}=\frac{(1,t)}{\sqrt{1+t^2}}, \tag{22.160}$$

dont la limite $t\to 0$ est bien le vecteur horizontal $(1,0)$.

La figure 22.7 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,

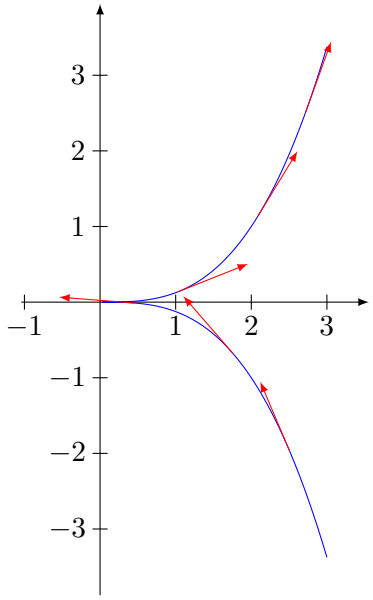


FIGURE 22.7 – Quelques tangentes de la courbe $\gamma(t) = (t^2, t^3)$.

ces vecteurs n'ont pas de limites quand $t \to 0$. Ce sont bien les *directions* qui ont une limite, parce que la direction ne tient pas compte du sens.

△

22.8 Repère de Frenet

Dans cette section, nous ne considérons que des courbes dans $\mathbb{R}^3$.

Proposition 22.54.
Soit $\gamma = (J, \gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^2$. Alors pour toute valeur de s dans J, nous avons

$$\tau(s) \cdot \tau'(s) = 0 \tag{22.161}$$

où $\tau(s) = \gamma'_N(s)$. C'est-à-dire que la dérivée seconde est perpendiculaire à la dérivée première.

Démonstration. La paramétrisation étant normale, nous avons

$$\|\gamma'_N(s)\|^2 = \sum_{i=1}^{n} x'_i(s)^2 = 1; \tag{22.162}$$

ce qui implique, en dérivant les deux membres, que

$$0 = 2\sum_{i=1}^{n} x'_i(s)x''_i(s), \tag{22.163}$$

c'est-à-dire exactement $\gamma'_N(s) \cdot \gamma''_N(s) = 0$; d'où la thèse. □

Remarque 22.55.
Si nous n'utilisons pas des coordonnées normales, la proposition 22.54 n'est pas spécialement vraie. Prenons par exemple la courbe qui donne la parabole :

$$\gamma(t) = (t, t^2) \tag{22.164a}$$
$$\gamma'(t) = (1, 2t) \tag{22.164b}$$
$$\gamma''(t) = (0, 2) \tag{22.164c}$$

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.

Définition 22.56.
Soit $\gamma=(J,\gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^2$.

*(1) Le **vecteur unitaire tangent** est donné par le corollaire 22.51 : $\tau(t)=\gamma'(t)/\|\gamma'(t)\|$.*

*(2) La **normale principale** est le vecteur $\tau'(s)$. Le **vecteur unitaire normal** est le vecteur*

$$\nu(s)=\frac{\tau'(s)}{\|\tau'(s)\|}=\frac{\gamma_N''(s)}{\|\gamma_N''(s)\|}. \tag{22.165}$$

Nous déduirons une formule plus pratique en dehors des coordonnées normales en (22.193).

*(3) La **courbure** au point $\gamma_N(s)$ est le réel*

$$c(s)=\|\tau'(s)\|=\|\gamma_N''(s)\|. \tag{22.166}$$

Note : il y a une notion de courbure signée qui sera donnée dans la définition 22.70.

*(4) Le **rayon de courbure** est le réel*

$$R(s)=\frac{1}{c(s)}=\frac{1}{\|\gamma_N''(s)\|}. \tag{22.167}$$

Par la proposition 22.54, nous avons $\nu(s)\cdot\tau(s)=0$. En combinant toutes les formules, nous avons les différentes expressions suivantes pour le vecteur normal unitaire :

$$\nu(s)=\frac{\gamma_N''(s)}{c(s)}=\frac{\tau'(s)}{\|\tau'(s)\|}=\frac{\tau'(s)}{c(s)}=R(s)\tau'(s)=R(s)\gamma_N''(s). \tag{22.168}$$

Proposition 22.57.
La fonction courbure s'écrit $c=\|\gamma_N'\times\gamma_N''\|$.

Démonstration. Par la proposition 11.36 nous avons :

$$\langle\gamma_N',\gamma_N''\rangle^2+\|\gamma_N'\times\gamma_N''\|^2=\|\gamma_N'\|^2\|\gamma_N''\|^2=\|\gamma_N''\|^2 \tag{22.169}$$

parce que, la paramétrisation étant normale, $\|\gamma_N'\|=1$. Mais $\langle\gamma_N',\gamma_N''\rangle=0$, donc il reste $\|\gamma_N'\times\gamma_N''\|^2=\|\gamma_N''\|^2$, d'où

$$c(s)=\|\gamma_N''(s)\|=\|\gamma_N'(s)\times\gamma_N''(s)\| \tag{22.170}$$

pour chaque s dans J. □

Définition 22.58.
*Soit s un point birégulier (c'est-à-dire $\gamma_N'(s)\neq0$ et $\gamma_N''(s)\neq0$) de l'arc normal $\gamma=(J,\gamma_N)$. Le **vecteur unitaire de la binormale** est le vecteur*

$$\beta(s)=\tau(s)\times\nu(s) \tag{22.171}$$

Par leurs définitions, τ et ν sont unitaires, tandis que la proposition 22.54 montre qu'ils sont également orthogonaux. Les propriétés du produit vectoriel font que β est également unitaire, et simultanément orthogonal à τ et à ν.

Définition 22.59.
*Le repère orthonormal $\{\gamma_N(s),\tau(s),\beta(s)\}$ est le **repère de Frenet** au point $\gamma_N(s)$.*

Lemme 22.60.
Le vecteur unitaire normal est donné par $\nu(s)=\beta(s)\times\tau(s)$.

Démonstration. Ceci est une application de la formule d'expulsion (11.71) et de l'orthonormalité de la base de Frenet :

$$\beta\times\tau=(\tau\times\nu)\times\tau=-(\nu\cdot\tau)\tau+(\tau\cdot\tau)\nu=\nu. \tag{22.172}$$

□

22.8.1 Torsion

Décomposons le vecteur $\beta'(s)$ dans la base de Frenet. Pour cela nous allons utiliser la proposition 11.290 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 β' le long de τ et β sont nulles. Le vecteur β' sera donc colinéaire à ν.

D'abord, étant donné que la norme de $\beta(s)$ est constante par rapport à s, nous avons

$$0 = \frac{d}{ds}\|\beta(s)\|^2 = 2\beta'(s) \cdot \beta(s). \tag{22.173}$$

Ensuite, nous dérivons la définition $\beta(s) = \tau(s) \times \nu(s)$ en utilisant la formule de Leibnitz (13.265) :

$$\beta'(s) = \tau'(s) \times \nu(s) + \tau(s) \times \nu'(s). \tag{22.174}$$

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 :

$$\beta'(s) = t(s)\nu(s). \tag{22.175}$$

Définition 22.61.
*Soit $\gamma = (J, \gamma_N)$ un arc paramétré normal de classe $\mathcal{C}^3$. La **torsion** de γ au point $\gamma_N(s)$ est le réel*

$$t(s) = \|\beta'(s)\| = \|\tau(s) \times \nu'(s)\|. \tag{22.176}$$

*Lorsque $t(s) \neq 0$, le réel $T(s) = \frac{1}{t(s)}$ est le **rayon de torsion** de γ 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

$$\begin{aligned} \tau'(s) &= c(s)\nu(s) \\ \beta'(s) &= t(s)\nu(s). \end{aligned} \tag{22.177}$$

Il reste à décomposer $\nu'(s)$. Définissons α_τ, α_ν et α_β (qui peuvent dépendre de s) par

$$\nu'(s) = \alpha_\tau \tau(s) + \alpha_\nu \nu(s) + \alpha_\beta \beta(s). \tag{22.178}$$

En vertu de la proposition 11.290, nous avons

$$\begin{aligned} \alpha_\tau &= \langle \nu'(s), \tau(s) \rangle = -\langle \nu(s), \tau'(s) \rangle = -\langle \nu(s), c(s)\nu(s) \rangle = -c(s), \\ \alpha_\nu &= \langle \nu'(s), \nu(s) \rangle = 0, \\ \alpha_\beta &= \langle \nu'(s), \beta(s) \rangle = -\langle \nu(s), \beta'(s) \rangle = -t(s), \end{aligned} \tag{22.179}$$

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** :

$$\begin{pmatrix} \tau'(s) \\ \nu'(s) \\ \beta'(s) \end{pmatrix} = \begin{pmatrix} 0 & c(s) & 0 \\ -c(s) & 0 & -t(s) \\ 0 & t(s) & 0 \end{pmatrix} \begin{pmatrix} \tau(s) \\ \nu(s) \\ \beta(s) \end{pmatrix}. \tag{22.180}$$

Proposition 22.62.
Si s est un point birégulier, alors la torsion est donnée par

$$t(s) = -\frac{(\gamma_N' \times \gamma_N'') \times \gamma_N'''}{\|\gamma_N'(s) \times \gamma_N''(s)\|^2}. \tag{22.181}$$

Démonstration. Par l'équation (22.168), nous avons $\gamma_N''(s) = c'(s)\nu(s)$, et par conséquent

$$\gamma_N'''(s) = c'(s)\nu(s) + c(s)\nu'(s) = c'(s)\nu(s) + c(s)\big[-c(s)\tau(s) - t(s)\beta(s)\big], \tag{22.182}$$

où nous avons utilisé la formule de Frenet pour $\nu'(s)$. Par ailleurs, sachant le corollaire 22.52 et la formule de Frenet pour τ', nous avons

$$\gamma_N' \times \gamma_N'' = \tau(s) \times \tau'(s) = \tau(s) \times c(s)\nu(s) = c(s)\beta(s). \tag{22.183}$$

En combinant les deux dernières équations, et en se souvenant que la base de Frenet et orthonormale,

$$(\gamma_N' \times \gamma_N'') \cdot \gamma_N'''(s) = -c(s)^2 t(s), \tag{22.184}$$

et donc, en remplaçant $c(s)$ par la formule (22.170),

$$t(s) = -\frac{(\gamma_N' \times \gamma_N'') \cdot \gamma_N'''}{\|\gamma_N' \times \gamma_N''\|^2}. \tag{22.185}$$

□

22.9 Hors des coordonnées normales

Remarque 22.63.
Notons que la définition de τ est donnée pour tout arc $\mathcal{C}^1$ régulier (I, γ) par $\tau(t) = \gamma'(t)/\|\gamma'(t)\|$. La propriété $\tau = \gamma_N'$ n'est valable que lorsque la paramétrisation est normale. Les autres définitions ont toutes été données dans le cas d'une paramétrisation normale.

La remarque 22.63 nous incite à exprimer toute la base de Frenet en terme de γ lorsque la paramétrisation n'est pas normale. Étant donné que nous pouvons toujours faire le changement de variable $\gamma(t) = \gamma_N\big(\phi(t)\big)$ (proposition 22.46), il est possible d'exprimer les vecteurs τ, ν et β ainsi que les réels c et t en fonction de γ 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 corollaire 22.52, τ est donné par $\tau = \gamma_N'$.

Vecteur unitaire normal Par la définition 22.56, ν est donné par $\nu = \frac{\tau'}{\|\tau'\|}$.

Vecteur unitaire de la binormale Par la définition 22.56, β est donné par $\beta = \tau \times \nu$.

Courbure Par la définition 22.56, c est donné par $c = \|\tau'\|$.

Torsion Par la définition 22.61, t est donné par $t = \|\beta'\|$.

Le schéma du changement de variable est

$$\begin{array}{ccc} t \in I & \xrightarrow{f} & \mathbb{R}^3 \\ {\scriptstyle\phi}\downarrow & \nearrow_{g} & \\ s \in J & & \end{array} \tag{22.186}$$

La difficulté ne sera pas d'éliminer γ_N de toutes les formules, mais bien de se débarrasser des fonctions ϕ qui arrivent quand nous exprimons γ_N en termes de γ, et en particulier lorsque nous voulons exprimer les dérivées de γ_N en termes de γ et de ses dérivées.

Regardons d'abord comment les dérivées de γ_N s'expriment en termes de γ. En utilisant le fait que $\gamma_N(s) = (\gamma \circ \phi^{-1})(s)$ et que $\|\gamma_N'(s)\| = 1$, nous avons

$$\gamma_N'(s) = \frac{\gamma_N'(s)}{\|\gamma_N'(s)\|} = \frac{(\gamma \circ \phi^{-1})'(s)}{\|(\gamma \circ \phi^{-1})'(s)\|} = \frac{\gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)}{\|\gamma'\big(\phi^{-1}(s)\big)\|\,|(\phi^{-1})'(s)|} = \frac{\gamma'(t)}{\|\gamma'(t)\|} \tag{22.187}$$

où nous avons utilisé le fait que ϕ^{-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 (22.187) :

$$\gamma_N''(s) = \frac{\gamma''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)}{\|\gamma'\big(\phi^{-1}(s)\big)\|} + \gamma'\big(\phi^{-1}(s)\big)\frac{d}{ds}\Big[\|\gamma'\big(\phi^{-1}(s)\big)\|\Big]. \tag{22.188}$$

Le petit calcul suivant va nous permettre de simplifier cette expression :

$$(\phi^{-1})'(s) = (\phi^{-1})'\big(\phi(t)\big) = \frac{1}{\phi'(t)} = \frac{1}{\|\gamma'(t)\|}. \tag{22.189}$$

Donc

$$\gamma_N''(s) = \frac{\gamma''(t)}{\|\gamma'(t)\|^2} + \gamma'(t)\frac{d}{ds}\Big[\|\gamma'(t)\|\Big] \tag{22.190}$$

où il est entendu que $t = \phi^{-1}(s)$. Avec cette expression, nous ne nous sommes pas encore débarrassés de la fonction ϕ, mais nous allons voir que cela nous sera suffisant.

Pour le vecteur unitaire tangent $\tau(s)$, nous avons donc immédiatement

$$\tau(s) = \gamma_N'(s) = \frac{\gamma'(t)}{\|\gamma'(t)\|}. \tag{22.191}$$

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 (22.191) contient encore la fonction ϕ.

Géométriquement, le vecteur $\tau(s)$ est le vecteur normal unitaire de la courbe au point $\gamma_N(s)$. En utilisant les relations du diagramme (22.186), nous avons en réalité $\gamma_N(s) = \gamma_N\big(\phi(t)\big) = \gamma(t)$. 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

$$\begin{aligned}\gamma_N'(s) \times \gamma_N''(s) &= \frac{\gamma'(t)}{\|\gamma'(t)\|} \times \left(\frac{\gamma''(t)}{\|\gamma'(t)\|^2} + \gamma'(t)\frac{d}{ds}\Big[\|\gamma'(t)\|\Big]\right) \\ &= \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t)\|^3}\end{aligned} \tag{22.192}$$

parce que le deuxième terme dans la parenthèse est un multiple de $\gamma'(t)$, de telle sorte à ce que son produit vectoriel avec $\gamma'(t)/\|\gamma'(t)\|$ soit nul. En prenant la norme,

$$c(s) = \frac{\|\gamma'(t) \times \gamma''(t)\|}{\|\gamma'(t)\|^3}. \tag{22.193}$$

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 (22.191) et (22.168),

$$\beta(s) = \tau(s) \times \nu(s) = \gamma_N'(s) \times \frac{\gamma_N''(s)}{c(s)}. \tag{22.194}$$

Les formules (22.192) pour le produit vectoriel et (22.193) pour la courbure donnent ensuite

$$\beta(s) = \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t)\|^3} \cdot \frac{1}{c(s)} = \frac{\gamma'(t) \times \gamma''(t)}{\|\gamma'(t) \times \gamma''(t)\|}. \tag{22.195}$$

Cela donne le vecteur unitaire binormal au point $\gamma(t)$ en terme de $\gamma'(t)$ et $\gamma''(t)$.

La torsion demande d'utiliser la dérivée troisième de γ_N. Nous avons

$$\begin{aligned}
\gamma_N'''(s) &= (\gamma\circ\phi^{-1})'''(s)\\
&= \Big(\gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)\Big)''\\
&= \Big(\gamma''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)^2+\gamma'\big(\phi^{-1}(s)\big)(\phi^{-1})''(s)\Big)'\\
&= \gamma'''\big(\phi^{-1}(s)\big)(\phi^{-1})'(s)^3+v\\
&= \frac{\gamma'''\big(\phi^{-1}(s)\big)}{\|\gamma'(t)\|^3}+v \qquad \text{par (22.189)}
\end{aligned} \tag{22.196}$$

où v est un élément de $\langle\gamma''\big(\phi^{-1}(s)\big),\gamma'\big(\phi^{-1}(s)\big)\rangle$. Le vecteur v est donc perpendiculaire à $\gamma'\times\gamma''$ et donc à $\gamma_N'\times\gamma_N''$ à cause de la relation (22.192) 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 γ_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 (22.192) en termes de γ,

$$\begin{aligned}
t(s) &= -\frac{(\gamma_N'\times\gamma_N'')\cdot\gamma_N'''}{\|\gamma_N'\times\gamma_N''\|^2}\\
&= -\frac{(\gamma_N'\times\gamma_N'')\cdot\gamma'''(t)}{\|\gamma_N'\times\gamma_N''\|^2\,\|\gamma'(t)\|^2}\\
&= -\frac{(\gamma'\times\gamma'')\cdot\gamma'''}{\|\gamma'\times\gamma''\|^2}.
\end{aligned} \tag{22.197}$$

Dans cette expression, il est sous-entendu que tous les γ_N sont fonctions de s et tous les γ sont fonction de t où s et t sont liés par $s=\phi(t)$.

Ce que nous avons prouvé est le

Théorème 22.64.
Pour tout représentant (I,γ), les éléments métriques (τ,ν,β,c,t) au point $\gamma(t)$ s'expriment en fonction de $\gamma(t)$, $\gamma'(t)$, $\gamma''(t)$ et $\gamma'''(t)$.

Lemme 22.65.
Si γ est le graphe de la fonction y alors la courbure de γ est donnée par la formule

$$c\big(\gamma(t)\big)=\frac{|y''(t)|}{\big(1+y'(t)^2\big)^{3/2}} \tag{22.198}$$

Démonstration. Nous avons :

$$\gamma(t)=\big(t,y(t)\big) \tag{22.199a}$$
$$\gamma'(t)=\big(1,y'(t)\big) \tag{22.199b}$$
$$\gamma''(t)=\big(0,y''(t)\big). \tag{22.199c}$$

Il s'agit maintenant seulement d'utiliser la formule (22.193) en se souvenant comment on calcule un produit vectoriel[7] :

$$\gamma'\times\gamma''=\begin{vmatrix}e_1 & e_2 & e_3\\ 1 & y' & 0\\ 0 & y'' & 0\end{vmatrix}=y''. \tag{22.200}$$

□

7. Définition 11.28.

22.10 Tracer des courbes paramétriques dans $\mathbb{R}^2$

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 $\gamma(t) = \big(x(t), y(t)\big)$, 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 γ 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 :

$$R(t) = \frac{\|\gamma(t)\|^3}{\|\gamma'(t) \times \gamma''(t)\|}. \tag{22.201}$$

En pratique, le produit vectoriel se calcule comme ceci :

$$\gamma'(t) \times \gamma''(t) = \begin{vmatrix} e_x & e_y & e_z \\ x'(t) & y'(t) & 0 \\ x''(t) & y''(t) & 0 \end{vmatrix} = (x'y'' - x''y')e_z. \tag{22.202}$$

Le centre du cercle osculateur va se trouver quelque part sur la normale. Le vecteur normal est donné par

$$n(t) = J\frac{\gamma'(t)}{\|\gamma'(t)\|} \tag{22.203}$$

où J est la rotation d'angle $\frac{\pi}{2}$:

$$J\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} y'(t) \\ -x'(t) \end{pmatrix}. \tag{22.204}$$

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.

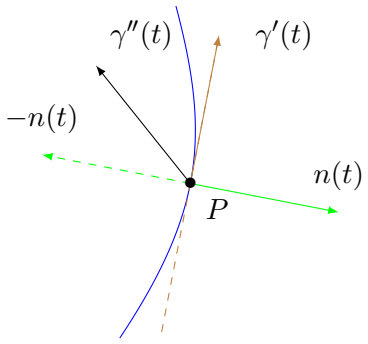


FIGURE 22.8 – De quel côté de $\gamma'(t)$ se trouvent $n(t)$ et $-n(t)$?

La difficulté maintenant est de savoir qui de $n(t)$ 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 la figure 22.8, 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 s'il est négatif, il faut choisir $n'(t)$.

Le truc pour obtenir le signe de $x'y'' - x''y'$ est de faire

$$\frac{(\gamma' \times \gamma'') \cdot e_z}{\|\gamma' \times \gamma''\|}. \tag{22.205}$$

Le centre de courbure sera donc situé à la position

$$\Omega(t) = \gamma(t) - n(t)\frac{\|\gamma(t)\|^3}{\|\gamma'(t) \times \gamma''(t)\|^2}(\gamma' \times \gamma'') \cdot e_z \tag{22.206}$$

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

$$\Omega_x(t) = x(t) - y'(t)\frac{x'^2 + y'^2}{x'y'' - x''y'} \tag{22.207a}$$

$$\Omega_y(t) = y(t) + x'(t)\frac{x'^2 + y'^2}{x'y'' - x''y'}. \tag{22.207b}$$

Quelques exemples de cercles osculateurs sont sur la figure 22.9.

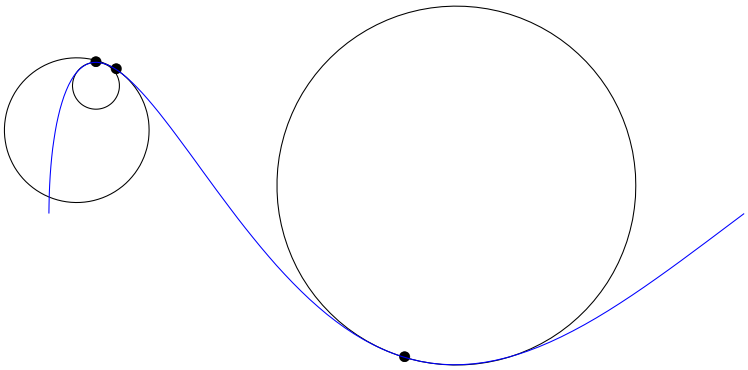


FIGURE 22.9 – Exemple de cercles osculateurs.

22.11 Courbes planes

Définition 22.66.
*Une courbe $\gamma\colon [a,b] \to \mathbb{R}^n$ est **fermée** si $\gamma(a) = \gamma(b)$. Elle est **simple** si $\gamma(t) \neq \gamma(t')$ dès que $t, t' \in]a,b[$ et $t \neq t'$.*

Définition 22.67.
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 $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ a cette régularité.

22.11.1 Angle

Lemme 22.68 ([317]).
Soient des courbes régulières γ et σ de classe C^2 de l'intervalle ouvert I vers $\mathbb{R}^2$. Soit $\theta_0 \in \mathbb{R}$ tel que

$$\frac{\gamma'(t_0) \cdot \sigma'(t_0)}{\|\gamma'(t_0)\|\|\sigma'(t_0)\|} = \cos(\theta_0) \tag{22.208a}$$

$$\frac{\gamma'(t_0) \cdot J\sigma'(t_0)}{\|\gamma'(t_0)\|\|\sigma'(t_0)\|} = \cos(\theta_0). \tag{22.208b}$$

Alors il existe une unique fonction différentiable $\theta\colon I \to \mathbb{R}$

$$\frac{\gamma'(t) \cdot \sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} = \cos\big(\theta(t)\big) \tag{22.209a}$$

$$\frac{\gamma'(t) \cdot J\sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} = \sin\big(\theta(t)\big). \tag{22.209b}$$

Démonstration. Il suffit de prendre

$$f(t) = \frac{\gamma'(t) \cdot \sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} \tag{22.210}$$

et

$$g(t) = \frac{\gamma'(t) \cdot J\sigma'(t)}{\|\gamma'(t)\|\|\sigma'(t)\|} \tag{22.211}$$

dans la proposition 19.40. Ces courbes sont de classe C^1 parce que γ et σ sont de classe C^2. □

22.11.2 Courbure signée

Nous avons déjà défini la courbure d'une courbe en la définition 22.56. Nous introduisons maintenant la courbure signée qui est propre à la dimension deux.

Définition 22.69.
*La **structure complexe** sur $\mathbb{R}^2$ est l'application*

$$\begin{aligned} J\colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x,y) &\mapsto (-y,x). \end{aligned} \tag{22.212}$$

Définition 22.70.
*La **courbure signée** de la courbe $\gamma\colon I \to \mathbb{R}^2$ (I est un intervalle dans $\mathbb{R}$) est la fonction*

$$\kappa(t) = \frac{\gamma''(t) \cdot J\gamma'(t)}{\|\gamma'(t)\|^3} \tag{22.213}$$

où J est la structure complexe de la définition 22.69.

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

$$v \times w = -(v \cdot Jw)e_3. \tag{22.214}$$

En particulier pour tout $v \in \mathbb{R}^2$ nous avons

$$v \cdot Jv = 0. \tag{22.215}$$

Lemme 22.71.
Soir une courbe régulière $\gamma\colon [a,b] \to \mathbb{R}^2$ et un difféomorphisme $h\colon [c,d] \to [a,b]$. Si nous posons $\sigma = \gamma \circ h$ alors

$$\kappa_\sigma(u) = \operatorname{sgn}\big(h'(u)\big)\kappa_\gamma\big(h(u)\big). \tag{22.216}$$

Démonstration. Nous utilisons la définition (22.213) de la courbure signée. La règle de dérivation en chaîne donne :

$$\sigma'(u) = \gamma'\big(h(u)\big)h'(u) \tag{22.217a}$$
$$\sigma''(u) = \gamma''\big(h(u)\big)h'(u)^2 + \gamma'\big(h(u)\big)h''(u). \tag{22.217b}$$

La numérateur de $\kappa_\sigma(u)$ est :

$$(\gamma'' \circ h)h'^2 \cdot J(\gamma' \circ h)h' + h''(\gamma' \circ h) \cdot J(\gamma' \cdot h)h' \tag{22.218}$$

dont le second terme est nul parce que $v \cdot Jv = 0$. Il nous reste donc

$$\kappa_\sigma(u) = \frac{(h')^3}{|h'|^3}\frac{(\gamma'' \circ h) \cdot J(\gamma' \circ h)}{\|\gamma' \circ h\|^3} = \operatorname{sgn}(h')\kappa_\gamma\big(h(u)\big). \tag{22.219}$$

□

Lemme 22.72 ([317]).
Si γ_N est un arc paramétré normal, alors

$$\gamma_N''(s) = \kappa(s)J\gamma_N'(s). \tag{22.220}$$

Démonstration. Vu que la paramétrisation est normale, $\gamma_N' \cdot \gamma_N' = 1$, et en dérivant, $\gamma_N'' \cdot \gamma_N' = 0$. Donc γ_N'' est un multiple de $J\gamma'$. En tenant compte du fait que la paramétrisation est normale, la courbure est

$$\kappa(s) = \gamma_N''(s) \cdot J\gamma'(s). \tag{22.221}$$

En y injectant $\gamma''(s) = \lambda(s)J\gamma'(s)$ nous trouvons

$$\kappa(s) = \lambda(s)J\gamma'(s) \cdot J\gamma'(s) = \lambda(s). \tag{22.222}$$

Donc le facteur de proportionnalité est $\kappa(s)$. □

Théorème 22.73.
Soit une courbe $\gamma\colon I \to \mathbb{R}^2$ de classe C^2.

(1) γ est une partie de droite si et seulement si $\kappa(t) = 0$ pour tout t.

(2) γ est une partie d'un cercle de rayon $r > 0$ si et seulement si $|\kappa(s)| = \frac{1}{r}$.

Démonstration. Si γ 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 utilisant une paramétrisation normale de γ, ce qui ne change pas que la courbure reste nulle. Nous avons par le lemme 22.72 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 γ est un cercle, la paramétrisation normale est

$$\gamma(t) = R\begin{pmatrix}\cos(\frac{t}{R})\\ \sin(\frac{t}{R})\end{pmatrix} \tag{22.223a}$$

$$\gamma'(t) = \begin{pmatrix}-\sin(\frac{t}{R})\\ \cos(\frac{t}{R})\end{pmatrix} \tag{22.223b}$$

$$\gamma''(t) = -\frac{1}{R}\begin{pmatrix}\cos(\frac{t}{R})\\ \sin(\frac{t}{R})\end{pmatrix}. \tag{22.223c}$$

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

$$\begin{aligned}\beta\colon [b, c] &\to \mathbb{R}^2\\ t &\mapsto \gamma(t) + rJ\gamma'(t).\end{aligned} \tag{22.224}$$

Nous avons $\beta'(t) = \gamma'(t) + rJ\gamma''(t)$. Mais par le lemme 22.72 nous avons $\gamma'' = kJ\gamma' = \frac{1}{r}J\gamma'$. Donc

$$\beta'(t) = \gamma'(t) - \gamma'(t) = 0. \tag{22.225}$$

Du coup β est constante : $\beta(t) = a$. Alors $a = \gamma(t) + rJ\gamma'(t)$ et en particulier

$$\|\gamma(t) - a\| = \|rJ\gamma'(t)\| = r. \tag{22.226}$$

Donc effectivement γ reste sur un cercle de rayon r et de centre a. □

Définition 22.74 ([317]).
*La **courbure totale** de la courbe $\gamma\colon [a, b] \to \mathbb{R}^2$ est le nombre*

$$K = \int_a^b \kappa(t)\|\gamma'(t)\|dt. \tag{22.227}$$

Lemme 22.75.
La courbure signée ne change pas sous reparamétrisation positive, et change de signe sous reparamétrisation négative.

Démonstration. Soit la courbe $\gamma\colon [a,b]\to\mathbb{R}^2$ et un difféomorphisme $h\colon [a,b]\to[c,d]$. Il s'agit d'intégrer la relation (22.216) en effectuant le changement de variables $t=h(u)$:

$$K_\gamma = \int_a^b \kappa_\gamma(t)\|\gamma'(t)\|dt \tag{22.228a}$$
$$= \int_c^d \underbrace{\kappa_\gamma\big(h(u)\big)}_{\kappa_\sigma(u)} \|\gamma'\big(h(u)\big)\|h'(u)du. \tag{22.228b}$$

En utilisant le fait que $\sigma'(u)=\gamma'\big(h(u)\big)h'(u)$ nous avons alors

$$K_\gamma = \int_c^d \kappa_\sigma(u)\left\|\frac{\sigma'(u)}{h'(u)}\right\|h'(u)du \tag{22.229a}$$
$$= \int_c^d \kappa_\sigma(u)\|\sigma'(u)\|\frac{h'(u)}{|h'(u)|}du \tag{22.229b}$$
$$= \operatorname{sgn}(h')K_\sigma. \tag{22.229c}$$

□

Lemme-définition 22.76 ([317]).
Soit une courbe régulière $\gamma\colon I\to\mathbb{R}^2$ de classe C^2 et t_0 dans l'intérieur de I. Soit $\theta_0\in\mathbb{R}$ tel que

$$\frac{\gamma'(t_0)}{\|\gamma'(t_0)\|} = \big(\cos(\theta_0),\sin(\theta_0)\big). \tag{22.230}$$

Alors il existe une unique fonction différentiable $\theta\colon I\to\mathbb{R}$ telle que $\theta(t_0)=\theta_0$ et

$$\frac{\gamma'(t)}{\|\gamma'(t)\|} = \begin{pmatrix}\cos\big(\theta(t)\big)\\ \sin\big(\theta(t)\big)\end{pmatrix} \tag{22.231}$$

pour tout $t\in I$.

*Cette fonction est l'**angle** de γ déterminé par θ_0.*

Démonstration. Soit $\beta(t)=(t,0)$; alors $\beta'(t)=(1,0)$ et nous avons

$$\gamma'\cdot\beta' = \gamma'_x \tag{22.232a}$$
$$\gamma'\cdot J\beta' = \gamma'_y. \tag{22.232b}$$

Par la proposition 22.68 il existe une unique fonction $\theta\colon I\to\mathbb{R}$ telle que $\theta(t_0)=\theta_0$ et

$$\begin{cases}\cos\big(\theta(t)\big) = \dfrac{\gamma'\cdot\beta'}{\|\gamma'\|\|\beta'\|} = \dfrac{\gamma'_x}{\|\gamma'\|} & (22.233a)\\[2ex] \sin\big(\theta(t)\big) = \dfrac{\gamma'\cdot J\beta'}{\|\gamma'\|\|\beta'\|} = \dfrac{\gamma'_y}{\|\gamma'\|}. & (22.233b)\end{cases}$$

Une telle fonction est bien celle que l'on demande ici. □

Lemme 22.77.
Si γ est une courbe régulière de classe C^2, alors sa courbure et son angle vérifient la relation

$$\theta'(t) = \|\gamma'(t)\|\kappa(t). \tag{22.234}$$

Démonstration. Par définition de l'angle (lemme 22.76) nous avons

$$\frac{\gamma'(t)}{\|\gamma'(t)\|} = \big(\cos\theta(t),\sin\theta(t)\big). \tag{22.235}$$

Dérivant cela,

$$\frac{\gamma''(t)}{\|\gamma'(t)\|}+\gamma'(t)\frac{d}{dt}\left(\frac{1}{\|\gamma'(t)\|}\right)=\theta'(t)\begin{pmatrix}-\sin\big(\theta(t)\big)\\ \cos\big(\theta(t)\big)\end{pmatrix} \tag{22.236a}$$

$$=\theta'(t)J\begin{pmatrix}\cos\big(\theta(t)\big)\\ \sin\big(\theta(t)\big)\end{pmatrix} \tag{22.236b}$$

$$=\theta'(t)\frac{J\gamma'(t)}{\|\gamma'(t)\|}. \tag{22.236c}$$

Nous prenons le produit scalaire de cette égalité avec $J\gamma'$ en tenant compte du fait que $\gamma'\cdot J\gamma'=0$:

$$\frac{\gamma''\cdot J\gamma'}{\|\gamma'\|}=\frac{\theta'}{\|\gamma'\|}J\gamma'\cdot J\gamma'. \tag{22.237}$$

En remarquant que $Jv\cdot Jv=\|v\|^2$ nous trouvons $\theta'\|\gamma'\|=\|\gamma'\|^2\kappa_\gamma$ et donc

$$\theta'(t)=\|\gamma'(t)\|\kappa_\gamma(t), \tag{22.238}$$

ce qu'il fallait prouver. □

Ce lemme nous fournit la formule attendue pour la courbure totale.

Lemme 22.78 ([317]).
Soit une courbe régulière de classe C^2 $\gamma\colon[a,b]\to\mathbb{R}^2$. Sa courbure totale est donnée par

$$K=\theta(b)-\theta(a). \tag{22.239}$$

Démonstration. Il suffit de remplacer dans la définition (22.227) de la courbure totale l'intégrante par son expression du lemme 22.77 :

$$K=\int_a^b\kappa(t)\|\gamma'(t)\|dt=\int_a^b\theta'(t)dt=\theta(b)-\theta(a) \tag{22.240}$$

par le théorème 15.233. □

22.11.3 Degré, indice et homotopie

Définition 22.79 ([317]).
*Le **nombre de tours** d'une courbe fermée de classe C^1 $\gamma\colon[a,b]\to\mathbb{R}^2$ est le nombre*

$$\mathrm{Turn}(\gamma)=\frac{1}{2\pi}\int_a^b\kappa(t)\|\gamma'(t)\|dt \tag{22.241}$$

où κ est la courbure signée définie en 22.70.

Lemme-définition 22.80.
*Soit une application continue $\phi\colon S^1\to S^1$. Un **relèvement** de σ est une application $\varphi\colon\mathbb{R}\to\mathbb{R}$ une application continue telle que*

$$\phi\big(\cos(t),\sin(t)\big)=\begin{pmatrix}\cos\big(\varphi(t)\big)\\ \sin\big(\varphi(t)\big)\end{pmatrix}. \tag{22.242}$$

*Le **degré** est l'entier $\deg(\phi)$ tel que*

$$\varphi(2\pi)-\varphi(0)=2\deg(\phi)\pi. \tag{22.243}$$

Ce nombre ne dépend pas du choix du relèvement φ.

Démonstration. Soient φ_1 et φ_2, deux applications qui satisfont les contraintes. Nous avons une fonction continue $n\colon S^1 \to \mathbb{R}$ telle que

$$\varphi_1(t) - \varphi_2(t) = 2\pi n(t). \tag{22.244}$$

La fonction n ne pouvant prendre que des valeurs entières et étant continue, elle est constante. Par conséquent

$$\varphi_1(2\pi) - \varphi_1(0) = \varphi_2(2\pi) - \varphi_2(0), \tag{22.245}$$

ce que nous voulions. □

Définition 22.81.
Soit une courbe fermée $\gamma\colon [0,L] \to \mathbb{R}^2$ de classe C^1. Nous posons

$$\Phi_\gamma(t) = \frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} \tag{22.246}$$

*où $\tilde{\gamma}(u) = \gamma\big(\frac{L}{2\pi}u\big)$. L'**indice de rotation** de γ est le degré de Φ_γ, c'est-à-dire*

$$\mathrm{Ind}(\gamma) = \deg(\Phi_\gamma). \tag{22.247}$$

Notons que dans cette définition, Φ_γ n'est rien d'autre que le vecteur unitaire tangent à γ, ramené à $[0,2\pi]$.

Proposition 22.82.
Pour une courbe fermée de classe[8] C^2, l'indice de rotation est égal au nombre de tours.

Démonstration. Soit $\gamma\colon [0,L] \to \mathbb{R}^2$ une courbe vérifiant les hypothèses. Par définition, $\mathrm{Ind}(\gamma) = \deg(\Phi_\gamma)$ où

$$\Phi_\gamma(t) = \frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} = \begin{pmatrix} \cos\tilde{\theta}(t) \\ \sin\tilde{\theta}(t) \end{pmatrix} \tag{22.248}$$

où nous avons noté $\tilde{\theta}$ l'angle tournant[9] de $\tilde{\gamma}$ et θ celui de γ. Vu que $\tilde{\theta}$ vérifie les hypothèses de la définition 22.80 nous pouvons calculer le degré de Φ_γ par

$$\deg(\Phi_\gamma) = \frac{1}{2\pi}\big(\tilde{\theta}(2\pi) - \tilde{\theta}(0)\big) = \frac{1}{2\pi}\int_0^{2\pi} \tilde{\theta}'(s)ds. \tag{22.249}$$

Il faut trouver le lien entre θ et $\tilde{\theta}$. Pour cela nous notons que

$$\frac{\tilde{\gamma}'(t)}{\|\tilde{\gamma}'(t)\|} = \frac{\gamma'\left(\frac{L}{2\pi}t\right)}{\|\gamma'\left(\frac{L}{2\pi}\right)\|}. \tag{22.250}$$

En comparant avec (22.248) il vient

$$\tilde{\theta}(t) = \theta\big(\frac{L}{2\pi}t\big) \tag{22.251}$$

et

$$\tilde{\theta}'(s) = \frac{L}{2\pi}\theta'\left(\frac{L}{2\pi}s\right). \tag{22.252}$$

Le changement de variables $t = \frac{L}{2\pi}s$ est donc tout vu dans l'intégrale (22.249) :

$$\deg(\Phi_\gamma) = \frac{1}{2\pi}\int_0^{2\pi} \frac{L}{2\pi}\theta'\big(\frac{L}{2\pi}s\big)ds = \frac{1}{2\pi}\int_0^{L} \theta'(t)dt = \frac{1}{2\pi}\int_0^L \kappa_\gamma(t)\|\gamma'(t)\|dt = \mathrm{Turn}(\gamma) \tag{22.253}$$

où nous avons aussi utilisé le lemme 22.77 qui donne le lien entre θ' et κ. □

8. 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.
9. Définition 22.76.

Définition 22.83 (homotopie de chemins fermés)**.**
Les courbes fermées $\gamma_0, \gamma_1 : [0, L] \to Y$ (Y est un espace topologique) sont ***homotopes*** *s'il existe une application continue*

$$F : [0, 1] \times [0, L] \to Y \tag{22.254}$$

telle que

(1) $F(0, t) = \gamma_0(t)$ *pour tout* t,

(2) $F(1, t) = \gamma_1(t)$ *pour tout* t,

(3) $F(u, L) = F(u, 0)$ *pour tout* u.

L'application F est l'homotopie entre γ_0 et γ_1.

Proposition 22.84 (Homotopie, degré et indice[317])**.**
Il y a deux résultats à ne pas confondre.

(1) Si $\phi_i : S^1 \to S^1$ *sont homotopes, alors* $\deg(\phi_1) = \deg(\phi_2)$.

(2) Si $\gamma_i : [0, L] \to \mathbb{R}^2$ *sont homotopes, alors* $\mathrm{Ind}(\phi_1) = \mathrm{Ind}(\phi_2)$.

Démonstration. Nous allons décomposer la preuve en deux parties.

Le degré pour les applications $S^1 \to S^1$ Soient deux applications homotopes $\gamma_i : S^1 \to S^1$. Si F est l'homotopie[10] entre γ_0 et γ_1, nous posons $\gamma_u(t) = F(u, t)$ qui est encore une courbe fermée $\gamma_u : S^1 \to S^2$. Nous pouvons donc considérer le degré de γ_u. C'est un entier n_u qui vérifie

$$\tilde{\gamma}_u(2\pi) - \tilde{\gamma}_u(0) = 2\pi n_u. \tag{22.255}$$

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.

Indice de rotation pour des courbes dans $\mathbb{R}^2$ Soient maintenant γ_0 et γ_1 deux courbes homotopes de $[0, L]$ dans $\mathbb{R}^2$. Par définition, $\mathrm{Ind}(\gamma_i) = \deg(\Phi_{\gamma_i})$. Prouvons alors que Φ_{γ_0} et homotope à Φ_{γ_1}. De cette façon, la première partie de la preuve conclura à

$$\mathrm{Ind}(\gamma_0) = \deg(\Phi_{\gamma_0}) = \deg(\Phi_{\gamma_1}) = \mathrm{Ind}(\gamma_1). \tag{22.256}$$

Nous savons que

$$\Phi_\gamma = \frac{\gamma'\left(\frac{L}{2\pi}t\right)}{\|\gamma'\left(\frac{L}{2\pi}t\right)\|}, \tag{22.257}$$

donc en posant

$$\Phi_F(u, t) = \frac{\gamma_u'\left(\frac{L}{2\pi}t\right)}{\|\gamma_u'\left(\frac{L}{2\pi}t\right)\|} \tag{22.258}$$

nous avons une homotopie entre Φ_{γ_0} et Φ_{γ_1}.

□

Théorème 22.85 ([317])**.**
Le nombre de tours d'une courbe simple fermée de classe C^2 dans $\mathbb{R}^2$ est ± 1.

Démonstration. Soit une telle courbe γ et son image Γ.

Choix d'un point et d'une tangente Si ℓ est une droite dans le plan, soit p le point de Γ le plus proche de ℓ. Il n'est peut-être pas unique, mais la parallèle à ℓ passant par p est une tangente à Γ telle que tout Γ se trouve d'un seul côté de ℓ_p. Pour voir cela, il suffit de choisir un système d'axes pour lequel ℓ est l'axe $y = 0$. La distance entre ℓ et les points de Γ est donnée par γ_y, et donc les extrema sont atteints là où $\gamma_y' = 0$, c'est-à-dire pour les points sur lesquels la tangente est parallèle à ℓ.

Étant donné qu'il y a (au moins) un maximum et un minimum distincts, pour pouvons choisir le point p de telle sorte que pour tout point $q \in \Gamma$, l'angle du vecteur $q - p$ soit entre α_0 et

10. Définition 22.83.

$\alpha_0 + \pi$ et non entre $\alpha_0 - pi$ et α_0. Ce choix revient à choisir p de telle sorte que Γ soit d'un côté ou de l'autre de ℓ_p.

Pour simplifier les notations plus tard nous choisissons ℓ_p horizontale, de telle sorte que $\alpha_0 = 0$, et que Γ est au dessus de ℓ_p. Les angles des vecteurs $q - p$ pour $q \in \Gamma$ sont donc tous entre 0 et π.

Définition de Σ Soit L la longueur[11] de γ et β, une paramétrisation de γ telle que $\beta(0) = p$. Nous considérons le triangle

$$\mathcal{T} = \{(t_1, t_2) \in \mathbb{R}^2 \text{ tel que } 0 \leqslant t_1 \leqslant t_2 \leqslant L\} \tag{22.259}$$

et l'application sécante $\Sigma\colon \mathcal{T} \to S^1$ définie par

$$\Sigma(t_1, t_2) = \begin{cases} \dfrac{\beta'(t)}{\|\beta'(t)\|} & \text{si } t_1 = t_2 = t \\ -\dfrac{\beta'(0)}{\|\beta'(0)\|} & \text{si } t_1 = 0 \text{ et } t_2 = L \\ \dfrac{\beta(t_2) - \beta(t_1)}{\|\beta(t_2) - \beta(t_1)\|} & \text{sinon.} \end{cases} \tag{22.260}$$

Continuité de Σ Nous devons prouver les limites[12] suivantes :

$$\lim_{\substack{(t_1,t_2)\to(t,t)\\ 0\leqslant t_1\leqslant t\leqslant t_2\leqslant L}} \frac{\beta(t_1) - \beta(t_2)}{\|\beta(t_1) - \beta(t_2)\|} = \frac{\beta'(t)}{\|\beta'(t)\|} \tag{22.261a}$$

$$\lim_{\substack{(t_1,t_2)\to(0,L)\\ 0\leqslant t_1\leqslant t\leqslant t_2\leqslant L}} \frac{\beta(t_1) - \beta(t_2)}{\|\beta(t_1) - \beta(t_2)\|} = -\frac{\beta'(0)}{\|\beta'(0)\|} \tag{22.261b}$$

Nous commençons par (22.261a) en multipliant et divisant par $t_2 - t_1$:

$$\frac{\beta(t_2) - \beta(t_1)}{\|\beta(t_2) - \beta(t_1)\|} = \frac{\beta(t_2) - \beta(t_1)}{t_2 - t_1} \frac{t_2 - t_1}{\|\beta(t_2) - \beta(t_1)\|}. \tag{22.262}$$

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 (13.950) :

$$\beta(t_2) = \beta(t_1) + (t_2 - t_1)\beta'(t_1) + \alpha(t_2 - t_1) \tag{22.263}$$

où α est une fonction ayant la propriété $\lim_{t\to 0} \frac{\alpha(t)}{t} = 0$. Nous avons à calculer la limite de

$$\frac{(t_1 - t_1)\beta'(t_1) + \alpha(t_2 - t_1)}{t_2 - t_1} = \beta'(t_1) + \frac{\alpha(t_2 - t_1)}{t_2 - t_1}. \tag{22.264}$$

Prendre la limite $(t_1, t_2) \to (t, t)$ donne bien $\beta'(t)$ parce que β 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 (22.261b).

La limite (22.261b). Nous prenons la prolongation périodique de β. 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) :

$$\beta(t_1) = \beta(-\epsilon) + (t_1 + \epsilon)\beta'(-\epsilon) + \alpha(t_1 + \epsilon). \tag{22.265}$$

Après multiplication et division par $t_1 + \epsilon$, la première limite à calculer est celle de

$$\frac{\beta(-\epsilon) - \beta(t_1)}{t_1 + \epsilon} = -\beta'(-\epsilon) + \frac{\alpha(t_1 + \epsilon)}{t_1 + \epsilon}, \tag{22.266}$$

11. Définition 22.4.
12. Limite au sens de la définition 7.62, en sachant qu'elle est unique par la proposition 7.65.

pour $(\epsilon, t_1) \to (0,0)$. Cela donne bien $-\beta'(0)$. La seconde limite à calculer est celle de

$$\frac{t_1+\epsilon}{\|\beta(-\epsilon)-\beta(t_1)\|} = \frac{t_1+\epsilon}{\|-(t_1+\epsilon)\beta'(-\epsilon)-\alpha(t_1+\epsilon)\|} = \frac{t_1+\epsilon}{\|(t_1+\epsilon)\beta'(-\epsilon)+\alpha(t_1+\epsilon)\|}. \tag{22.267}$$

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

$$\big|a-|b|\big| \leqslant |a+b| \leqslant \big|a+|b|\big|. \tag{22.268}$$

Nous avons l'encadrement

$$\frac{\big|(t_1+\epsilon)\beta'(-\epsilon)-|\alpha(t_1+\epsilon)|\big|}{t_1+\epsilon} \leqslant \frac{\|(t_1+\epsilon)\beta'(-\epsilon)+\alpha(t_1+\epsilon)\|}{t_1+\epsilon} \tag{22.269a}$$

$$\leqslant \frac{\big|(t_1+\epsilon)\beta'(-\epsilon)+|\alpha(t_1+\epsilon)|\big|}{t_1+\epsilon} v \tag{22.269b}$$

Les limites des deux extrêmes existent et valent $\beta'(0)$; la règle de l'étau 13.157 conclu.

Deux chemins homotopes Nous considérons dans $\mathcal{T}$ les points $A=(0,0)$, $B=(0,L)$ et $C=(L,L)$. Nous allons considérer les chemins direct de A à C et celui passant via B. Et comme ces chemins doivent être paramétrés de 0 à 2π, il faut faire un peu attention. Nous définissons les chemins

$$\sigma_i : S^1 \to S^1 \tag{22.270}$$

par

$$\sigma_1(t) = \Sigma\left(\frac{L}{2\pi}t, \frac{L}{2\pi}t\right) = \frac{\beta'(tL/2\pi)}{\|\beta'(tL/2\pi)\|} = \frac{\tilde{\beta}'(t)}{\|\tilde{\beta}'(t)\|} = \Phi_\beta \tag{22.271a}$$

$$\sigma_2(t) = \begin{cases} \Sigma(0, \frac{L}{\pi}t) & \text{si } 0 \leqslant t \leqslant \pi \\ \Sigma\big(\frac{L}{\pi}(t-\pi), L\big) & \text{si } \pi \leqslant t \leqslant 2\pi \end{cases} \tag{22.271b}$$

où nous avons repris les notations de la définition 22.81. Notons que pour σ_2, en $t=\pi$ les deux expressions donnent

$$\sigma_2(\pi) = \Sigma(0,L) = -\frac{\beta'(0)}{\|\beta'(0)\|}. \tag{22.272}$$

Ce sont des chemins fermés parce que

$$\sigma_1(2\pi) = \Sigma(L,L) = \frac{\beta'(L)}{\|\beta'(L)\|} = \frac{\beta'(0)}{\|\beta'(0)\|} = \Sigma(0,0) = \sigma_1(0). \tag{22.273}$$

Notons que dans toutes ces définitions et calculs, nous avons utilisé de façon assez cruciale la définition 22.67 pour définir la dérivée de β en $t=0$.

Les chemins σ_1 et σ_2 sont homotopes par construction.

Indices et degrés La proposition 22.84(1) nous donne $\deg(\sigma_1) = \deg(\sigma_2)$. Nous avons alors la chaîne d'égalités

$$\deg(\sigma_2) = \deg(\sigma_1) = \deg(\Phi_\beta) = \mathrm{Ind}(\beta) = \mathrm{Ind}(\gamma) = \mathrm{Turn}(\gamma) \tag{22.274}$$

où les justifications sont :

(1) $\deg(\sigma_2) = \deg(\sigma_1)$ par homotopie : proposition 22.84.

(2) $\deg(\sigma_1) = \deg(\Phi_\beta)$ parce que $\sigma_1 = \Phi_\beta$.

(3) $\deg(\Phi_\beta = \mathrm{Ind}(\beta)$ par définition 22.81 de l'indice.

(4) $\mathrm{Ind}(\beta) = \mathrm{Ind}(\gamma)$ par invariance de l'indice sous reparamétrisation.

(5) $\mathrm{Ind}(\gamma) = \mathrm{Turn}(\gamma)$ par la proposition 22.82.

Il nous reste à montrer que $\deg(\sigma_2) = \pm 1$.

La géométrie de σ_2 Notons que par définition les valeurs de $\sigma_2(t)$ sont les vecteurs (unitaires) joignant p aux points de Γ lorsque $0 \leqslant t \leqslant \pi$, et les vecteurs inverses pour $\pi \leqslant t \leqslant 2\pi$. Plus précisément nous avons, si $0 < a < \pi$:

$$\sigma_2(a) = \Sigma(0, \frac{L}{\pi}a) = \frac{\beta(\frac{L}{\pi}a) - \beta(L)}{\|\beta(\frac{L}{\pi}a) - \beta(L)\|} \tag{22.275}$$

et

$$\sigma_2(\pi + a) = \Sigma(\frac{L}{\pi}a, L) = \frac{\beta(L) - \beta(\frac{L}{\pi}a)}{\|\beta(L) - \beta(\frac{L}{\pi}a)\|}. \tag{22.276}$$

En sachant que $\beta(0) = \beta(L)$ nous avons alors

$$\sigma_2(\pi + a) = -\sigma_2(a). \tag{22.277}$$

Notons aussi que le fait que γ soit une courbe simple assure que le numérateur et le dénominateur de σ_2 ne s'annulent pas autrement que pour $t = L$ ou $t = 0$.

Un relèvement pour σ_2 Ces propriétés motivent cette idée pour le relèvement de σ_2 :

$$\begin{aligned} \theta\colon [0, 2\pi[&\to [0, 2\pi[\\ t &\mapsto \text{l'angle du vecteur } \sigma_2(t), \end{aligned} \tag{22.278}$$

avec $\theta(2\pi)$ définit 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 θ petit à petit.

Étant donné que ℓ_p est tangente à Γ au point p, la droite ℓ_p est parallèle à $\beta'(0)$, et l'angle entre ℓ_p et $\beta'(0)$ est soit 0 soit π. Donc $\theta(0)$ devrait valoir soit 0 soit π.

Nous commençons par définir ceci :

$$\begin{aligned} \theta\colon [0, \pi[&\to [0, \pi] \\ t &\mapsto \text{angle de } \sigma_2(t) = \arccos\big(\sigma_2(t)_x\big). \end{aligned} \tag{22.279}$$

C'est le choix d'avoir ℓ_p horizontale et Γ au dessus de ℓ_p qui nous assure que pour tout $t \in [0, \pi]$, l'angle de $\sigma_2(t)$ peut être choisi entre 0 et π. De plus cette fonction est continue en tant que partie de la fonction $\arccos\colon [-1, 1] \to [0, \pi]$ qui est elle-même continue par la proposition 7.77.

Le fait que θ soit continue est assuré par le fait que σ_2 est C^∞.

Si $\theta(0) = 0$ Cela correspond à la situation des vecteurs rouges sur la figure 22.10.

Vu que $\sigma_2(\pi) = -\sigma_2(0)$, l'angle de σ_2 en $t = \pi$ est le supplémentaire de celui en $t = 0$. Mais pour $0 \leqslant t < \pi$, $\theta(t)$ prend ses valeurs entre 0 et π, 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 $0 \leqslant t \leqslant \pi$. Pour assurer la continuité en $t = pi$ nous sélectionnons la place $]\pi, 2\pi]$; et nous définissons

$$\begin{aligned} \theta\colon]\pi, 2\pi[&\to]\pi, 2\pi] \\ t &\mapsto \text{angle de } \sigma_2(t). \end{aligned} \tag{22.280}$$

Le nombre $\theta(2\pi)$ est défini par continuité. Il doit valoir $\theta(2\pi) = 2\pi$.

La fonction θ ainsi définie est un relèvement pour σ_2, et le degré peut être calculé :

$$\deg(\sigma_2) = \frac{1}{2\pi}\big(\theta(2\pi) - \theta(0)\big) = 1. \tag{22.281}$$

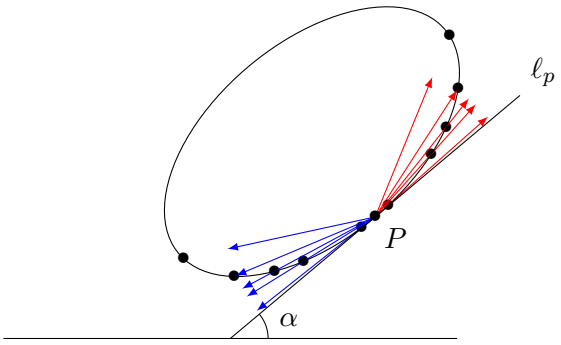


FIGURE 22.10 – Les vecteurs représenant σ_2 dans le cas où $\beta'(0)$ est dans le sens de ℓ_p ou dans le sens inverse. Pour le sport nous avons dessiné la situation avec une droite ℓ quelconque plutôt que horizontale.

Si $\theta(0) = \pi$ Cela correspond à la situation des vecteurs bleus sur la figure 22.10.

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

$$\begin{aligned}\theta\colon]\pi, 2\pi[&\to]-\pi, 0]\\ t &\mapsto \text{angle de } \sigma_2(t).\end{aligned} \tag{22.282}$$

Par continuité nous devons avoir $\theta(2\pi) = \theta(\pi) - \pi$ et donc $\theta(2\pi) = -\pi$.

Le degré de σ_2 est alors

$$\deg(\sigma_2) = \frac{1}{2\pi}\big(\theta(2\pi) - \theta(0)\big) = -1. \tag{22.283}$$

Conclusion Nous avons montré que le degré de σ_2 est 1 ou -1, et en remontant les égalités (22.274) nous déduisons que $\text{Turn}(\gamma) = \pm 1$.

□

22.12 Courbes fermées planes

22.12.1 Cercle circonscrit

La proposition suivante est dans le même esprit que l'ellipse de John-Loewer [13].

Proposition-définition 22.86 (Cercle circonscrit[368, 1])**.**
Soit une courbe fermée simple et continue $\gamma\colon [0,1] \to \mathbb{R}^2$. Soit Γ son image. Il existe un unique cercle de rayon minimum contenant Γ. Ce cercle est le ***cercle circonscrit*** *à γ.*

Il a les propriétés suivantes :

(1) Le cercle circonscrit à γ coupe Γ en au moins deux points distincts.

(2) Tout arc du cercle circonscrit plus grand que le demi-cercle intersection Γ.

Démonstration. Division de la preuve.

Existence L'application γ étant continue, l'ensemble Γ est compact (théorème 7.86). Nous considérons l'ensemble Q des formes quadratiques de la forme

$$q_{a,r}(x) = \|a - x\|^2 - r^2 \tag{22.284}$$

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

$$q\big(\gamma(t)\big) \leqslant 0 \tag{22.285}$$

13. Proposition 18.117.

pour tout $t \in [0,1]$.

Cet ensemble A est non vide parce que Γ est compact et donc borné ; il existe donc une boule qui contient Γ en son intérieur.

Montrons que A est fermé dans Q. Si $q_{a,r} \notin A$ alors il existe t_0 tel que $q_{a,r}\big(\gamma(t_0)\big) > 0$. Par continuité, il existe un voisinage de (a,r) dans $\mathbb{R}^3$ et donc de $q_{a,r}$ dans Q tel que $q_{a',r'}\big(\gamma(t_0)\big)$ reste strictement positif pour tout (a',r') dans ce voisinage.

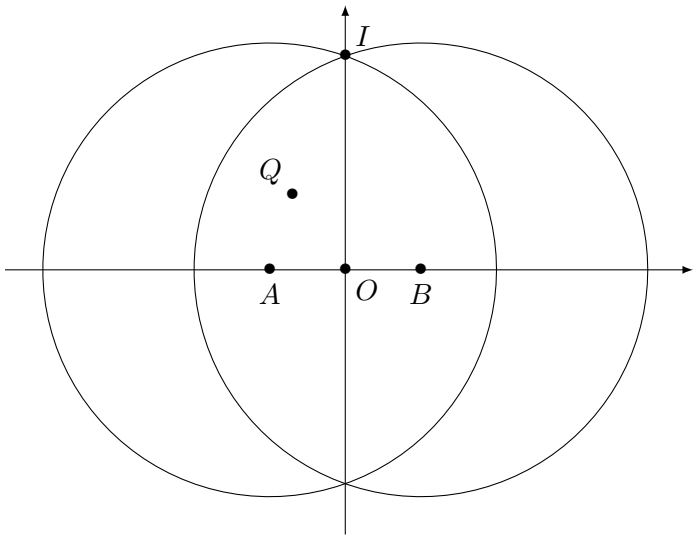
En particulier l'ensemble

$$\{r \in \mathbb{R}^+ \text{ tel que } \exists\,(a,r) \text{ tel que } q_{a,r} \in A\} \tag{22.286}$$

est fermé et borné vers le base. De plus $r = 0$ n'est pas dans cet ensemble. Il possède donc un minimum strictement positif. Le cercle correspondant donne l'existence.

Unicité En ce qui concerne l'unicité, si Γ est contenu dans les boules $B(a,R)$ et $B(b,R)$ alors

$$\Gamma \subset B(a,R) \cap B(b,R). \tag{22.287}$$



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

$$(B_x - Q_x)^2 + Q_y^2 < R. \tag{22.288}$$

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 $Q_x^2 + Q_y^2 \leqslant I_y^2$. Il s'agit simplement de calculer

$$Q_x^2 + Q_y^2 \leqslant Q_x^2 + R^2 - B_x^2 + 2B_xQ_x - Q_x^2 = R^2 - B_x^2 + 2B_xQ_x \leqslant R^2 - B_x^2 = I_y^2 \tag{22.289}$$

parce que $B_x > 0$ et $Q_x < 0$.

Nous concluons que Γ est inclus dans un cercle de rayon plus petit que R et donc que R n'est pas minimum. D'où l'unicité.

Au moins deux intersections Nous nommons C le cercle circonscrit à γ, et nous écrivons, pour $p \in \Gamma$

$$r(p) = d(p,C) \tag{22.290}$$

la distance entre p et C. Cela est une fonction continue sur le compact Γ. Elle atteint donc ses bornes sur Γ.

Si $C \cap \Gamma = \varnothing$ alors $r(p) > 0$ pour tout p et le minimum est également strictement positif. Soit r_0 ce minimum. Alors le cercle C_2 même centre que C mais de rayon $r_0/2$ n'intersecte pas non plus Γ parce que

$$d(p,C) \leqslant d(p,C_2) + d(C_2,C) \tag{22.291}$$

où $d(p,C)=r(p)$ et $d(C_2,C)=r_0/2$.

$$d(p,C_2)\geqslant r(p)-\frac{r_0}{2}\geqslant\frac{r_0}{2}>0. \tag{22.292}$$

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. 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 $p=(p_x,p_y)$,

$$r(p)=\sqrt{1-p_y^2}+p_x. \tag{22.293}$$

Cela est encore une fonction continue sur Γ qui atteint son minimum valant r_0. Alors le cercle de centre $(\frac{r_0}{2},0)$ et de même rayon contient encore Γ mais n'a plus de points d'intersection avec Γ.

Enfin, tout nombre de points d'intersection entre C et Γ 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.

Intersection avec les demi-arcs Supposons, en fixant encore les axes, que le cercle circonscrit soir encore centré en $(0,0)$ et que Γ 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 Γ mais ne l'intersectant pas.

□

22.12.2 Description locale

Définition 22.87.
Une courbe plane différentiable est ***convexe*** *si son graphe est en tout point d'un seul côté de sa tangente.*

Lemme 22.88 ([1])**.**
Soit une courbe simple, convexe $\gamma\colon[0,1]\to\mathbb{R}^2$ que nous supposons être de classe C^1. Nous notons Γ l'image de γ. Alors Γ est localement le graphe d'une fonction convexe (définition 18.76).

Démonstration. Soit $p\in\Gamma$ et l_p la tangente à Γ 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 Γ se trouve dans la partie $y>0$. De plus nous paramétrons γ de telle sorte à avoir $\gamma(0)=p=(0,0)$.

Vu que $l_p\equiv y=0$ est la tangente à Γ nous avons $\gamma_y'=0$ et $\gamma_x'(0)>0$. Nous en déduisons que $\gamma_x'(t)>0$ pour tout $t\in B(0,\delta)$ pour δ suffisamment petit. Nous posons alors

$$g(x)=\gamma_y\big(\gamma_x^{-1}(x)\big) \tag{22.294}$$

qui est bien définie parce que γ_x est une bijection entre $B(0,\delta)$ et son image. La fonction g est continue par la proposition 7.78 et même dérivable par la proposition 13.120. De plus, vu la formule (13.274), la fonction g^{-1} est de classe C^1 parce que $(g^{-1})'$ est une composée d'applications continues.

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 $g''\geqslant 0$ partout. Cela prouve que g est convexe par la caractérisation 18.82.

Si g est seulement de classe C^1 Le graphe de g correspond au graphe de Γ. Nous montrons que g est convexe en utilisant la caractérisation de la proposition 18.87.

La tangente au graphe de g en $x=x_0$, que nous notons l_0, est la tangente à Γ en $t=\gamma_x^{-1}(x_0)$. Le graphe de g, qui est une partie de Γ se trouve donc d'un seule côté de l_0.

Nous nous restreignons g à un compact I et nous considérons la fonction

$$d_a(x)=g(x)-l_a(x) \tag{22.295}$$

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 :

$$\begin{cases} d_0(x) \geqslant 0 & (22.296a) \\ d_{x_1}(x) \leqslant 0. & (22.296b) \end{cases}$$

Nous posons

$$s(a) = \sup_{x \in I} d_a(x), \tag{22.297}$$

qui est une fonction continue par la proposition 13.31. Vu le définitions, $s(0) \geqslant 0$ et $s(x_2) \leqslant 0$. Il existe donc $m \in [0, m_2]$ 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).

□

22.12.3 Enveloppe convexe

Proposition 22.89 ([369, 1])**.**
Soit une courbe simple, fermée et convexe $\gamma\colon [0,1] \to \mathbb{R}^2$ que nous supposons être de classe C^2. Nous notons Γ l'image de γ. Alors il existe un convexe D tel que $\partial D = \Gamma$.

Démonstration. Pour $p \in \Gamma$ nous notons l_p la tangente à Γ en p et H_p le demi-plan (fermé) contenant Γ. Nous posons

$$D = \bigcap_{p \in \Gamma} H_p. \tag{22.298}$$

Cet ensemble est convexe comme intersection de convexes et fermé comme intersection de fermés. 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 $p \in \partial D$.

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

$$r(q) = \frac{d(p, H_q)}{2} \tag{22.299}$$

définie sur Γ. C'est une fonction continue [14] strictement positive définie sur le compact Γ qui possède donc un minimum strictement positif. Si r_0 est ce minimum, alors $B(p, r_0)$ est inclue 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 $p \in l_q$.

Le point où ça décolle Nous supposons que $p \notin \Gamma$, sinon ce serait trop facile. Nous paramétrons γ de telle sorte à avoir $q = \gamma(0)$ et nous posons

$$T = \{t \in \mathbb{R}^+ \text{ tel que } l_{\gamma(t)} = l_q\}. \tag{22.300}$$

Cela est un fermé dans $\mathbb{R}$ parce que γ est de classe C^1. Nous posons $t_0 = \inf(T^c)$ et pour tout ϵ suffisamment petit,

$$\begin{cases} l_{\gamma(t_0+\epsilon)} \neq l_q, & (22.301a) \\ l_{\gamma(t_0-\epsilon)} = l_q. & (22.301b) \end{cases}$$

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 γ' nous aurions $l_{r'} \neq l_q$ pour tout $r' \in \gamma\big(B(t_0, \delta)\big)$.

14. L'équation de la droite l_q a des coefficients continus parce que γ est de classe C^1.

Graphe d'une fonction strictement convexe En suivant le lemme 22.88, l'ensemble Γ est localement (autour de r) le graphe d'une fonction convexe au-dessus de l_r. Soit $g\colon B(0,\delta)\to\mathbb{R}^2$ cette fonction convexe.

Soit $\epsilon>0$. Si $g''(x)=0$ sur $]0,\epsilon]$ alors g' y est constante. Mais $g'(0)=$, 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 Γ). 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 δ et ϵ de telle sorte que p_x soit plus grand à la fois que ϵ et δ.

Il existe donc un intervalle $[a,b]$ avec $a,b\geqslant 0$ et $a,b<p_x$ sur lequel g est strictement convexe.

La tangente qui tue En particulier $g(b)>0$ et le théorème des accroissements finis 13.129(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 $(b,f(b))$. Cette tangente, que nous nommons l_m, est en dessous de la corde, par strict 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 le points de coordonnées $(x,0)$ avec $x>m$ ne sont pas dans D. Pas de chance, p est un point de ce type.

La contradiction Nous avons prouvé que si $p\in\partial D\backslash\Gamma$ 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 22.90.
Bien que cela puisse paraitre évident dès le début, nous ne démontrerons que dans la proposition 22.92 que D est l'enveloppe convexe de Γ.

Corollaire 22.91.
Si $p\in\mathrm{Int}(D)$ alors toute droite passant par p intersecte Γ en exactement 2 points.

Démonstration. Vu que la partie D est bornée, toute droite passant par son intérieur coupe ∂D en au moins deux points (un dans chaque sens, et en utilisant le lemme de passage de douane 8.56).

Soit ℓ une droite passant par p et supposons qu'elle coupe Γ 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 $p\in\mathrm{Int}(D)$, $q_1\in\partial D$ et $q_2\in\partial D$.

Vu que tout Γ 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 ℓ_{q_1}, parce que si $q_2\in\ell_{q_1}$ alors $\ell_{q_1}=\ell$, ce qui est impossible parce que ℓ passe par $p\in\mathrm{Int}(D)$.

Or p et q_2 sont de part et d'autres de ℓ_{q_1}, ce qui est impossible parce que Γ est d'un seul côté de cette droite. □

Proposition 22.92.
Soit D l'ensemble définit en (22.298).

(1) $D=\mathcal{C}\mathrm{onv}(\Gamma)$ *($\mathcal{C}\mathrm{onv}(\Gamma)$ désigne l'enveloppe convexe de Γ)*

(2) $\partial\,\mathcal{C}\mathrm{onv}(\Gamma)=\Gamma$.

Pour la définition d'enveloppe convexe, voir la définition 10.35.

Démonstration. L'ensemble D est un convexe contenant Γ. Donc $\mathcal{C}\mathrm{onv}(\Gamma)\subset D$. L'inclusion inverse est à prouver.

Soit $x\in D$. Si $x\in\partial D$ alors $x\in\Gamma$ (proposition 22.89) et donc $x\in\mathcal{C}\mathrm{onv}(\Gamma)$. Nous ne devons donc traiter que le cas $x\in\mathrm{Int}(D)$.

Le corollaire 22.91 nous dit que toute droite passant par x coupe Γ en exactement deux points. Soient p et q ces points. Alors $p,q\in\Gamma\subset\mathcal{C}\mathrm{onv}(\Gamma)$. Vu que $\mathcal{C}\mathrm{onv}(\Gamma)$ est convexe, tout le segment $[p,q]$ est dans $\mathcal{C}\mathrm{onv}(\Gamma)$, et en particulier $p\in\mathcal{C}\mathrm{onv}(\Gamma)$.

Nous passons à la seconde affirmation. Nous savons que $D = \mathcal{C}\text{onv}(\Gamma)$. En prenant le bord des deux côtés, $\partial D = \partial\,\mathcal{C}\text{onv}(\Gamma)$, donc $\Gamma = \partial\,\mathcal{C}\text{onv}(\Gamma)$. □

Lemme 22.93 (Des tangentes parallèles[370])**.**
Une courbe fermée γ de classe C^1 est convexe si et seulement si elle ne possède pas 3 tangentes parallèles distinctes.

Démonstration. Si le graphe Γ possédait trois tangentes parallèles distinctes, une serait entre les deux autres et le graphe Γ serait de part et d'autres de cette tangente. Dans ce cas, γ n'est pas convexe.

Nous montrons maintenant que si γ n'est pas convexe, alors elle possède trois tangentes parallèles distinctes. Pour $p \in \Gamma$ tel que Γ soit des deux côtés de l_p.

Soient Γ_1 et Γ_2 les parties de Γ délimitées par l_p. Nous notons q_i le point de Γ_i le plus éloigné de la droite l_p, il existe parce que G est compact et que la fonction distance à l_p est continue sur Γ. Les points p, q_1 et q_2 sont distincts (sinon l_p ne couperait pas Γ 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 extrema 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 extrema peut ne pas vérifier cette condition s'il n'est pas à l'intérieur du domaine. Dans notre cas, nous avons $\gamma\colon [0,1] \to \mathbb{R}^2$ qui est fermée : $\gamma(0) = \gamma(1)$. Donc en réalité nous pouvons considérer $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ et tous les points du domaine sont intérieurs au domaine. L'extremum doit donc vérifier la condition d'annulation de la dérivée. □

Exemple 22.94
Si la courbe n'est pas fermée, alors le lemme 22.93 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 directions). △

Lemme 22.95.
Soit γ une courbe convexe et ℓ une droite qui intersecte Γ mais qui n'est pas tangente. Alors l'intersection entre ℓ et Γ comprend au maximum 2 points.

Démonstration. Supposons que ℓ coupe Γ en trois points p, q, r (dans cet ordre). Vu que ℓ n'est pas tangente à Γ, la tangente ℓ_q est distincte de ℓ (et intersecte ℓ en l'unique point q). Les points p et r sont dans Γ et sont pourtant de deux côtés différents de ℓ_q. Contradiction avec la convexité de γ. □

La proposition suivante nous dit que si deux points de Γ ont la même tangente, alors entre ces deux points, Γ est le segment de droite les joignant.

Proposition 22.96 ([1])**.**
Soit $\gamma\colon [0, L] \to \mathbb{R}^2$ 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

$$\gamma\big([s_p, s_q]\big) = [p, q] \tag{22.302}$$

soit

$$\gamma\big([s_q, L]\big) = [p, q] \tag{22.303}$$

Démonstration. Nous considérons un système d'axe dans lequel $p = (0,0)$, $\ell_p = \ell_q \equiv y = 0$ et tel que $\gamma_y(t) \geqslant 0$ pour tout t. Nous choisissons enfin une paramétrisation de γ telle que $p = \gamma(0)$.

Si $\gamma_y\big([0, s_q]\big) = \{0\}$ alors par le théorème des valeurs intermédiaires 13.50, tous les points du type $(t, 0)$ avec $0 \leqslant t \leqslant q_x$ sont atteints par $\gamma(t)$ avec $0 \leqslant t \leqslant s_q$. De plus aucun autre point ne peut être atteint parce que γ étant simple, elle ne peut pas faire de retour en arrière.

Nous supposons donc que γ_y n'est pas identiquement nulle sur $[0,s_q]$; il existe donc un s_M avec $0<s_M<q_q$ qui maximise γ_y sur $[0,s_q]$. La tangente à Γ en $\gamma(s_M)$ est horizontale. Les droites $y=0$ et $y=\gamma_y(s_M)$ sont donc deux tangentes parallèles à Γ. Par le lemme des tangentes parallèles 22.93, 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 [15] à Γ.

Par le lemme 22.95, la droite $y=\frac{1}{n}$ ne peut intersecter Γ qu'en seulement deux points. Or le théorème des valeurs intermédiaires appliqué à γ_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

$$0<a_n<s_M \tag{22.304a}$$
$$s_M<b_n<s_q \tag{22.304b}$$

et $\gamma_y(a_n)=\gamma_y(b_n)=\frac{1}{n}$. Donc la droite $y=\frac{1}{n}$ intersecte γ deux fois dans $[0,s_q]$. En conséquence de quoi $\gamma_y(t)<\frac{1}{n}$ pour tout $t\in[q_q,L]$. Cela étant valable pour tout n nous avons $\gamma_y\big([s_q,L]\big)=\{0\}$ et nous sommes ramenés essentiellement au premier cas. □

Proposition 22.97.
Soit une courbe fermée simple γ de classe C^1 et une droite ℓ. Alors il existe au moins deux points distincts q_1,q_2 tels que $\ell_{q_1}\parallel\ell_{q_2}\parallel\ell$ avec $\ell_p\neq\ell_q$.

Démonstration. Pour $p\in\Gamma$ nous considérons le nombre $r(p)=d(p,\ell)$. En tant que fonction continue sur un compact, elle possède un minimum et un maximum. Dans un système d'axe pour lequel $\ell\equiv y=0$, la fonction r s'écrit $r(p)=p_y$ et les extrema arrivent en $\gamma(s)$ avec $\gamma_y'(s)=0$, ce qui signifie que les tangentes aux extrema sont parallèles à ℓ. 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. □

Corollaire 22.98.
Soit une courbe convexe fermée simple γ de classe C^2. L'ensemble $\mathcal{C}\mathrm{onv}(\Gamma)$ est compact.

Démonstration. Nous savons que $\mathcal{C}\mathrm{onv}(\Gamma)$ n'est autre que D par la proposition 22.92. Nous savons déjà que D est fermé. Il nous suffit donc de prouver qu'il est borné (théorème de Borel-Lebesgue 8.9). Nous considérons deux droites perpendiculaires et les 4 tangentes correspondantes par la proposition 22.97. Vu que Γ est d'un seul côté de chacune de ces tangentes, elle est contenue dans le rectangle délimité par ces 4 droites. □

22.12.4 Courbure et convexité

Lemme 22.99.
Soit $\gamma\colon[0,L]\to\mathbb{R}^2$ une courbe simple, fermée en paramétrisation normale. Alors l'application

$$\begin{aligned}\sigma\colon[0,L]&\to S^1\\ s&\mapsto\gamma(s)\end{aligned} \tag{22.305}$$

est surjective.

Démonstration. Le théorème 22.85 nous dit que si $\theta(0)=a$ alors $\theta(2\pi)=a+2\pi$ ou $a-2\pi$. Le théorème des valeurs intermédiaires nous dit alors que θ prend toutes les valeurs entre a et $a+2\pi$ ou $a-2\pi$. □

Proposition 22.100 ([370])**.**
Une courbe fermée simple de classe C^2 est convexe si et seulement si sa courbure est soit toujours positive, soit toujours négative.

15. À part $n=0$ et si par manque de chance, $\gamma_y(s_M)$ est un nombre de la forme $1/n$.

Démonstration. Nous considérons la courbe γ munie d'un paramétrage de vitesse 1, c'est-à-dire avec $\|\gamma'(t)\| = 1$ pour tout t. Si θ est sa fonction d'angle, alors nous avons $\theta' = \kappa$ par le lemme 22.77. Donc la fonction θ est monotone si et seulement si la courbure ne change pas de signe. Nous allons montrer que θ est monotone si et seulement si γ est convexe.

$\Rightarrow$ Nous supposons que θ est monotone et γ non convexe. Soient $p \in \Gamma$ tel que Γ soit des deux côtés de ℓ_p, et soient les points q_1, q_2 donnés par le lemme des tangentes parallèles 22.93 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

$$\theta(s_1) = \theta(s_2) + 2n\pi \tag{22.306}$$

pour un certain n. Mais θ est monotone et la différence entre sa première et sa dernière valeur doit valoir 2π ou -2π par le théorème 22.85. Donc n ne peut valoir que -1, 0 et 1.

Si $n = 0$ alors θ est constante sur $[s_1, s_2]$. Si $n = 1$ alors $\theta(s_1) = \sigma(s_2) + 2\pi$ alors que sur toute la courbe, θ ne peut faire que 2π. Donc θ est constant sur $[0, s_1]$ et sur $[s_2, L]$ (où L est le bord de la paramétrisation). Si $n = -1$, même conclusion.

Dans tous les cas, Γ 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 γ est convexe.

$\Leftarrow$ Nous supposons que γ est convexe, mais que θ n'est pas monotone. Il existe donc $s_1 < s_0 < s_2$ tels que

$$\theta(s_1) = \theta(s_2) \neq \theta(s_0). \tag{22.307}$$

Et vu le lemme 22.99, 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 22.93 est alors formelle, γ étant convexe, deux de ces tangentes doivent être identiques.

La proposition 22.96 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 θ est constante.

- La partie $\gamma\big([s_1, s_2]\big)$ ne peut pas être droite parce que nous avions 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 $\gamma\big([s_1, s_3]\big)$ ne peut pas être droite parce que $\theta(s_3) \neq \theta(s_1)$.
- La partie $\gamma\big([s_2, s_3]\big)$ ne peut pas être droite parce que $\theta(s_2) \neq \theta(s_1)$.

Nous sommes donc devant une contradiction.

Nous en concluons que θ doit être monotone.

□

22.12.5 Théorème des quatre sommets

Lemme 22.101.
Soit une droite ℓ 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 γ est de la forme $z = \gamma(t) = p + tq$. Nous avons :

$$\big(\gamma(t) - a\big) \cdot c = \big(\gamma(t) - p\big) \cdot Jq = tq \cdot Jq = 0. \tag{22.308}$$

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 Γ. □

Définition 22.102.
*Un **sommet** d'une courbe est un point d'extremum local de la courbure.*

Théorème 22.103 (Théorème des quatre sommets[43, 317])**.**
Soit un arc paramétrique $\gamma\colon \mathbb{R} \to \mathbb{R}^2$ fermé, simple et convexe[16] *de classe C^3 et T-périodique.*

Alors γ possède au moins 4 points critiques sur chaque période.

Démonstration. Nous supposons que la paramétrisation de γ soit normale.

Si la courbure κ est constante sur une partie ouverte de la (paramétrisation de la) courbe, alors tous les points de cette partie sont des sommets et le théorème est fait. Nous supposons que κ n'est pas constante et en particulier que Γ ne contient ni bouts de droites ni bouts de cercles (théorème 22.73).

La fonction κ étant de classe C^1 sur le compact Γ, elle admet au moins un maximum et un minimum distincts. Vu que ces points sont intérieurs, ils correspondent au changement de signe de κ'. Soient p et q ces points. Pour la simplicité nous supposons que γ 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 κ'.

Soit ℓ la droite passant par p et q. Tous les points du segment $[p, q]$ (qui sont dans $\mathcal{C}\text{onv}(\Gamma)$) ne peuvent pas être sur Γ (sinon nous aurions un morceau de droite). Donc certains points sont dans l'intérieur de $\mathcal{C}\text{onv}(\Gamma)$. Donc la droite ℓ passe par l'intérieur de $\mathcal{C}\text{onv}(\Gamma)$ et le corollaire 22.91 nous dit que la droite ℓ ne coupe Γ en seulement deux points.

Par conséquent, les ensembles $\gamma\big([0, s_q]\big)$ et $\gamma\big([s_q, T]\big)$ sont de part et d'autre de Γ. Vu qu'en ces points, κ' change de signe et qu'il ne change de signe en aucun autre points, la fonction κ' est positive d'un côté de ℓ et négative de l'autre.

D'autre part par le lemme 22.101, 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 ℓ et négative de l'autre.

En résumé les fonctions

$$s \mapsto \kappa'(s) \tag{22.309a}$$
$$s \mapsto \big(\gamma(s) - a\big) \cdot c \tag{22.309b}$$

changent de signe en même temps et le produit a donc un signe constant. Ce produit n'est de plus pas nul parce que κ' n'est nul sur aucun intervalle (sinon κ y serait constant et Γ un segment de droite) et $\big(\gamma(s) - a\big) \cdot c$ ne s'annule pour aucun s sauf ceux qui correspondent à p et q.

Nous avons donc

$$0 \neq \int_0^T \kappa'(s)\big(\gamma(s) - a\big) \cdot c\, ds \tag{22.310a}$$
$$= \underbrace{\Big[\big(\gamma(s) - a\big) \cdot c\kappa(s)\Big]_0^T}_{A=0} - \int_0^T \kappa(s)\gamma'(s) \cdot c\, ds \tag{22.310b}$$
$$= -\int_0^T \kappa(s)\big(\gamma'(s) \cdot c\big) ds \tag{22.310c}$$
$$= \int_0^T J\gamma''(s) \cdot c\, ds \tag{22.310d}$$
$$= J\int_0^T \gamma''(s) \cdot c\, ds \tag{22.310e}$$
$$= J\big[\gamma'(s) \cdot c\big]_0^T \tag{22.310f}$$
$$= 0. \tag{22.310g}$$

Justifications :

16. Par la proposition 22.100 nous pouvons aussi bien demander à la courbure d'être toujours strictement positive, comme le fait [43].

— L'expression A est nulle parce que les valeurs en 0 et en T sont identiques.

— Nous utilisons le lemme 22.72 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 κ'. Vu que la courbe est périodique, il en faut un nombre pair et donc un quatrième. □

L'exemple de l'ellipse montre qu'il n'y a pas lieu de chercher d'autres extrema de κ à part les 4 déjà trouvés.

Exemple 22.104
Nous trouvons les sommets de l'ellipse.

$$\gamma(t) = \big(a\cos(t), b\sin(t)\big) \tag{22.311a}$$
$$\gamma'(t) = \big(-a\sin(t), b\cos(t)\big) \tag{22.311b}$$
$$\gamma''(t) = -\big(a\cos(t), b\sin(t)\big) \tag{22.311c}$$
$$\tag{22.311d}$$

La courbure est

$$\kappa(t) = \frac{\gamma'' \cdot J\gamma'}{\|\gamma'\|^3} \tag{22.312a}$$
$$= \frac{-1}{[a^2\sin^2(t) + b^2\cos^2(t)]^{3/2}} \begin{pmatrix} a\cos(t) \\ b\sin(t) \end{pmatrix} \cdot \begin{pmatrix} -b\cos(t) \\ -a\sin(t) \end{pmatrix} \tag{22.312b}$$
$$= \frac{ab}{[a^2\sin^2(t) + b^2\cos^2(t)]^{3/2}}. \tag{22.312c}$$

Vu que $ab > 0$, les extrema de cela sont ceux du dénominateur et il suffit donc d'étudier les extrema de

$$f(t) = a^2\sin^2(t) + b^2\cos^2(t). \tag{22.313}$$

Nous avons

$$f'(t) = 2(a^2 - b^2)\cos(t)\sin(t), \tag{22.314}$$

fonction qui s'annule effectivement 4 fois sur une période. Deux maxima et deux minima. △

22.12.6 Le théorème de Jordan

Définition 22.105 ([371]).
*Une **courbe de Jordan** est une courbe simple fermée dans le plan.*

Le théorème suivant a un énoncé relativement simple, mais la démonstration est en réalité très longue.

Théorème 22.106 (Théorème de Jordan[371]).
Le complémentaire d'une courbe de Jordan Γ dans un plan affine réel est formé de exactement deux composantes connexes distinctes, dont l'une est bornée et l'autre non. Toutes deux ont pour frontière la courbe Γ.

Chapitre 23

Géométrie hyperbolique

23.1 Inversion

Lemme-définition 23.1.
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

$$\|AP\|\|AP'\| \tag{23.1}$$

*ne dépend pas du choix de la droite et est nommé la **puissance** du point A par rapport au cercle C.*

23.1.1 Cercles perpendiculaires

Définition 23.2.
*Deux cercles sont **perpendiculaires** lorsque leurs tangentes aux points d'intersection sont perpendiculaires.*

Lemme 23.3 ([1])**.**
Soient deux cercles perpendiculaires C_1 et C_2. Alors

(1) le centre de C_1 est hors de C_2.

(2) Si ℓ est une droite passant par ce le centre de C_1 (nommé O) et si ℓ 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 ℓ_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 $Q' = \ell_Q \cap \ell_{Q'}$.

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 P et P' sont sur le cercle C_2, donc tous les points du segment $[PP']$ 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 $[PP']$, ce qui prouve que P et P' ne sont pas de part et d'autre de O sur la droite (OPP'). □

Proposition 23.4 ([1])**.**
Soit un cercle C_1 de centre O et de rayon R.

(1) Un cercle C_2 est perpendiculaire à C_1 si et seulement s'il existe une droite ℓ passant par O telle que les points d'intersection $\{P, P'\} = \ell \cap C_2$ soient situés du même côté de O et vérifient

$$\|OP\|\|OP'\| = R^2. \tag{23.2}$$

(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 23.3(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 ℓ :

$$\|OP\|\|OP'\|. \tag{23.3}$$

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 :

$$\|OP\|\|OP'\| = R^2 \tag{23.4}$$

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 23.1) et nous avons

$$\|OQ\|\|OQ'\| = R^2, \tag{23.5}$$

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. □

23.1.2 Inversion

Proposition-définition 23.5 ([313]).
Soit un cercle C de centre O dans $\mathbb{R}^2$. Il existe une unique application

$$i_C : \mathbb{R}^2 \backslash O \to \mathbb{R}^2 \backslash O \tag{23.6}$$

telle que

(1) *$i_C(x) = x$ pour tout $x \in C$*

(2) *i_C échange l'intérieur et l'extérieur de C.*

(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 ℓ la droite (OP) et nous considérons un cercle C_2 passant par P et perpendiculaire à C (existence par la proposition 23.4). Nous avons $\ell \perp C$ (parce que ℓ 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

$$i_C(P) \in \ell \cap C_2. \tag{23.7}$$

Mais ℓ 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. □

23.6.
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 ℓ 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 23.7.
Lorsque nous disons qu'une inversion « conserve les droites passant par O », il y a pour sous-entendu que nous considérons la droites privée du point O, parce que de toutes façons, l'inversion n'est pas définie sur O.

Nous allons résoudre cet intéressant problème en 24.6.1, en ajoutant le point ∞ à toutes les droites.

Corollaire 23.8.
L'inversion i_C est une involution ($i_C^2 = \mathrm{Id}$).

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 23.6 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. □

Soit C_1 le cercle dans $\mathbb{C}$ de centre 0 et de rayon 1. Nous notons $\alpha\colon \mathbb{C} \to \mathbb{C}$ la dilatation de rapport R.

Proposition 23.9 ([1])**.**
Soit C le cercle de rayon R centré en 0 ($C = \alpha(C_1)$). Alors

$$i_{\alpha(C_1)} \circ \alpha = \alpha \circ i_{C_1}. \tag{23.8}$$

Démonstration. Soit $z = re^{i\theta}$; nous devons prouver que

$$i_C\big(\alpha(z)\big) = \alpha\big(i_{C_1}(z)\big). \tag{23.9}$$

Nous avons :

$$i_C\big(\alpha(z)\big) = i_C(Rre^{i\theta}) = R^2\frac{1}{Rr}e^{i\theta} = R\frac{1}{r}e^{i\theta} = \alpha\big(\frac{1}{r}e^{i\theta}\big) = \alpha\big(i_{C_1}(z)\big). \tag{23.10}$$

□

La définition 24.67 pour l'inversion sur $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ sera basée sur la proposition suivante.

Proposition 23.10 ([1]).
Soit le cercle de rayon R et de centre $a \in \mathbb{C}$. Alors

$$i_C(z) = \frac{R^2}{\bar{z} - \bar{a}} + a. \tag{23.11}$$

De plus si t_a est la translation de vecteur a, nous avons la décomposition

$$i_{t_a(C_R)} = t_a \circ i_{C_R} \circ t_{-a} \tag{23.12}$$

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

$$\|a - z'\|\|a - z\| = R^2, \tag{23.13}$$

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

$$\lambda\|a - z\|^2 = R^2 \tag{23.14}$$

et alors

$$\lambda = \frac{R^2}{(a - z)(\bar{a} - \bar{z})}. \tag{23.15}$$

En récrivant $a - z' = \alpha(a - z)$ avec cette valeur de λ nous trouvons

$$a - z' = \frac{R^2}{\overline{a - z}}, \tag{23.16}$$

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 23.9. □

Avant d'aller plus loin, donnons l'équation d'un cercle dans $\mathbb{C}$. Si C est un cercle de entre ω 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

$$(z - \omega)(\bar{z} - \bar{\omega}) = r^2. \tag{23.17}$$

Proposition 23.11 ([313, 372]).
Soit un cercle C de centre O. L'inversion

$$i_C \colon \mathbb{R}^2 \backslash \{0\} \to \mathbb{R}^2 \backslash \{0\} \tag{23.18}$$

transforme

(1) les droites passant par O sur 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 23.5, 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 ℓ perpendiculaire à (OP) et passant par P'.

Nous montrons maintenant que $i_C(C_1\backslash\{O\}) = \ell$. Soit $Q \in C_1\backslash\{P, O\}$. Nous posons

$$Q' = (QO) \cap \ell. \tag{23.19}$$

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 ℓ 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 [1] :

$$\frac{OP}{OQ} = \frac{OQ'}{OP'}. \tag{23.20}$$

Donc

$$\|OP\|\|OP'\| = \|OQ\|\|OQ'\|, \tag{23.21}$$

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

$$\|OQ\|\|OQ'\| = R^2, \tag{23.22}$$

c'est-à-dire que Q' est l'image de Q par i_C, et donc que

$$i_C\big(C_1\backslash\{O\}\big) \subset \ell. \tag{23.23}$$

Pour avoir l'inclusion inverse, il faut remarquer que ℓ 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 (23.23) à l'envers, il est loisible de prouver que $i_C(Q') = Q$ et donc que ℓ est bien inclue 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 ℓ ne passant pas par O.

— la droite d, perpendiculaire à ℓ 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 ℓ est l'image du cercle C_1. Or si $\|OP\| = r$ alors

$$\|OP'\| = \frac{R^2}{r}. \tag{23.24}$$

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.

(4) Pour cette partie, nous allons utiliser un peu de géométrie analytique dans $\mathbb{C}$ [2].

1. Faites bien le dessin : ce n'est pas une situation de Thalès ultra-standard de collège.
2. Principalement parce que je ne comprends pas le raisonnement fait dans [313].

Le cas centré Nous supposons que C est centré en 0 et de rayon, 1, ce telle sorte que $i_C(z) = \frac{1}{\bar{z}}$. Soit C_1 un cercle de centre ω et de rayon r, ne passant pas par 0, en particulier tel que $|\omega| \neq r$.

Si $z \in i(C_1)$ alors $i(z) \in C_1$ et nous avons

$$\big(i(z) - \omega\big)\overline{\big(i(z) - \omega\big)} = r^2. \tag{23.25}$$

En développant et en multipliant par $z\bar{z}$ nous trouvons

$$\bar{\omega}z + \omega\bar{z} + z\bar{z}(r^2 - \omega\bar{\omega}) = 1. \tag{23.26}$$

Nous pouvons diviser par $(r^2 - |\omega|^2)$ parce que C_1 ne passe pas par 0. En remettant en ordre, et en notant $s = r^2 - \omega\bar{\omega}$ pour plus de clarté,

$$\big(z - \frac{\omega}{s}\big)(\bar{z} - \frac{\bar{\omega}}{s}) - \frac{\omega\bar{\omega}}{s^2} = \frac{1}{s^2}, \tag{23.27}$$

ou encore

$$\big(z - \frac{\omega}{s}\big)(\bar{z} - \frac{\bar{\omega}}{s}) = \frac{1}{s} + \frac{\omega\bar{\omega}}{s}. \tag{23.28}$$

Pour que cela soit l'équation du cercle de centre $\frac{\omega}{s}$ et de rayon $\sqrt{\frac{1}{s} + \frac{\omega\bar{\omega}}{s^2}}$, il faut vérifier que

$$\frac{1}{s} + \frac{\omega\bar{\omega}}{s^2} \geqslant 0. \tag{23.29}$$

En multipliant par s^2, il s'agit de vérifier que $s + \omega\bar{\omega} \geqslant 0$, ce qui est correct parce que $s + \omega\bar{\omega} = r$.

En résumé, si $z \in i(C_1)$ alors z est dans le cercle C_2 de centre $\frac{\omega}{s}$ et de rayon $\frac{r}{s}$. Étant donné que $r \neq \omega$ nous savons que ce dernier cercle ne passe pas par 0.

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.

Le cas de rayon non unité Si C est un cercle quelconque, nous écrivons l'inversion du cercle C via la formule (23.11). Si $z \in i_C(C_1)$ alors $i(z) \in C_1$ et nous pouvons écrire

$$\big(i(z) - a\big)\big(\overline{i(z)} - \bar{a}\big) = R^2. \tag{23.30}$$

De là il faut déduire que z est sur un cercle ne passant pas par 0. Bons calculs...

□

Chapitre 24

Espaces projectifs

Sur les espaces projectifs : [373].

Définition 24.1.
*Soit E un espace vectoriel de dimension finie sur le corps commutatif $\mathbb{K}$. Nous définissons sur $E\backslash\{0\}$ 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 $P(E)$.*

Définition 24.2.
*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 $P(\mathbb{K}^{n+1})$.

Exemple 24.3
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

$$P_1(\mathbb{R}) = \{(1,0)\} \cup \{(x,1) \text{ tel que } x \in \mathbb{R}\}. \tag{24.1}$$

Le point $(1, 0)$ est dit « point à l'infini ». △

24.1 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.

Proposition 24.4.
Si F et G sont des sous-espaces vectoriels de E, alors

$$P(F) \cap P(G) = P(F \cap G) \tag{24.2}$$

et nous avons

$$\dim P(F) + \dim P(G) = \dim P(F+G) + \dim P(F \cap G). \tag{24.3}$$

Démonstration. Nous avons

$$P(F) = \{[v] \text{ tel que } v \in F\} \tag{24.4}$$

où les crochets signifient la classe par rapport à la relation de colinéarité. Nous avons alors

$$P(F) \cap P(G) = \{[v] \text{ tel que } v \in F \cap G\} = P(F \cap G). \tag{24.5}$$

Cela prouve le premier point.

En ce qui concerne l'équation (24.3), en considérant $\dim P(E) = \dim E - 1$ nous devons prouver l'égalité

$$\dim F + \dim G = \dim(F + G) + \dim(F \cap G) \tag{24.6}$$

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

$$\dim F + \dim G = 2\operatorname{Card}(B_1) + \operatorname{Card}(B_2) + \operatorname{Card}(b_3) \tag{24.7a}$$
$$\dim(F + G) = \operatorname{Card}(B_1) + \operatorname{Card}(b_2) + \operatorname{Card}(B_3) \tag{24.7b}$$
$$\dim(F \cap G) = \operatorname{Card}(B_1). \tag{24.7c}$$

De là la relation (24.3) se déduit immédiatement. □

Théorème 24.5 (incidence)**.**
Soient F et F deux sous-espaces vectoriels de E tels que

$$\dim P(F) + \dim P(G) \geqslant \dim P(E). \tag{24.8}$$

Alors $P(F) \cap P(G) \neq \varnothing$.

Démonstration. En utilisant les hypothèses et la proposition 24.4 nous avons

$$\dim P(E) + \dim P(G) = \dim P(F + G) + \dim P(F \cap G) \geqslant \dim P(E). \tag{24.9}$$

En passant aux espaces vectoriels correspondants,

$$\dim(F + G) + \dim(F \cap G) \geqslant \dim(E) + 1. \tag{24.10}$$

Mais nous avons aussi $\dim(F + G) \leqslant \dim(E)$ et par conséquent $\dim(F \cap G) \geqslant 1$. Au final, $\dim P(F \cap G) \geqslant 0$. 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). □

Exemple 24.6
Soient les plans $\Pi_1 \equiv x = 0$ et $\Pi_2 \equiv y = 0$. Nous avons

$$P(\Pi_1) = \{[0, y, 1]\} \cup \{[0, 1, 0]\} \tag{24.11a}$$
$$P(\Pi_2) = \{[x, 0, 1]\} \cup \{[1, 0, 0]\} \tag{24.11b}$$

où le crochet signifie la classe pour la colinéarité. Ces deux droites projectives ont comme point d'intersection le point $[0, 0, 1]$. △

Définition 24.7.
*Un **hyperplan projectif** est un sous-espace projectif de $P(E)$ de la forme $P(V)$ où V est un hyperplan de E.*

Définition 24.8.
*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 une plan vectoriel $D \subset E$.*

*Nous disons que trois points de $P(E)$ sont **alignés** lorsqu'il existe une droite projective les contenant.*

24.9.
Dans la définition 24.8 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 paraitre idiot, mais ce qui serait surement idiot est de l'utiliser sans s'en rendre compte.

Lemme 24.10.
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 $A = B = \pi(a)$. □

Lemme 24.11.
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.

Sens direct Soient A, B, C distincts et alignés dans $P(E)$. Alors il existe un plan vectoriel D tel que $A, B, D \in P(D)$.
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 λa 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.

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 $P(E)$.

□

Proposition 24.12.
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 (24.3) pour notre cas :

$$\underbrace{\dim d}_{=1} + \underbrace{\dim H}_{=n-2} = \underbrace{\dim P(D + V)}_{=n-1} + \dim P(D \cap V). \tag{24.12}$$

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. □

24.2 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

$$\begin{aligned} \phi\colon d \cup \{\infty\} &\to P(E) \\ (x, 1) &\mapsto \text{la droite vectorielle passant par } (x, 1) \\ \infty &\mapsto \text{la droite vectorielle passant par } (1, 0). \end{aligned} \tag{24.13}$$

Lemme 24.13.
Si nous munissons l'ensemble $d \cup \{\infty\}$ de la topologie compactifiée d'Alexandrov[1]*, la bijection* (24.13) *est un homéomorphisme.*

Soient maintenant les plans affines dans l'espace vectoriel E de dimension 3

$$\Pi_1 \equiv z = 0 \tag{24.14a}$$
$$\Pi_2 \equiv z = 1. \tag{24.14b}$$

Une droite (vectorielle) de E coupe Π_2 en un et un seul point, sauf si elle est contenue dans Π_1. Nous avons donc une bijection

$$\begin{aligned}\phi\colon P(E) &\to \Pi_2 \cup P(\Pi_1)\\ d &\mapsto \begin{cases}\Pi_2 \cap d & \text{si cette intersection est non vide}\\ d & \text{sinon.}\end{cases}\end{aligned} \tag{24.15}$$

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 (Π_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

$$\mathcal{E} \cup P(E) \to P(F). \tag{24.16}$$

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 24.14
Nous considérons les plans affines

$$\Pi_1 \equiv z = 0 \tag{24.17a}$$
$$\Pi_2 \equiv z = 1 \tag{24.17b}$$

et nous avons la bijection

$$P(E) = \Pi_2 \cup P(\Pi_1). \tag{24.18}$$

Un plan affine D a deux possibilités : soit il coupe Π_2 en une droite, soit il est égal à Π_1. Si $D \cap \Pi_2 = d$ (d est une droite affine), alors nous avons

$$P(D) = d \cup \{\infty_D\}, \tag{24.19}$$

ce qui justifie la terminologie comme quoi $P(D)$ est une droite dans $P(E)$. △

Soit E un espace vectoriel de dimension 3 et le plan projectif $P(E)$. Nous avons deux types de droites projectives :

(1) D'abord nous avons la droite à l'infini, donnée[2] 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.

Exemple 24.15
Étudions un peu le second type de droites. D'abord si deux droites sont parallèles, leurs points à

1. Définition 7.61.
2. Dans notre représentation usuelle du plan projectif $z = 1$.

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

$$\begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 2 \\ y \\ 1 \end{pmatrix}. \tag{24.20}$$

En normalisant, ce sont les vecteurs

$$\frac{1}{\sqrt{2+y^2}} \begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} \quad \text{et} \quad \frac{1}{\sqrt{5+y^2}} \begin{pmatrix} 2 \\ y \\ 1 \end{pmatrix}, \tag{24.21}$$

et toutes deux tendent vers le vecteur $(0, 1, 0)$ pour $y \to \infty$. △

Lemme 24.16.
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 24.15.

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 a contient les vecteurs $(1, y, 0)$ et le point à l'infini $(0, 1, 0)$. La droite affine d' a pour équation paramétriques

$$\begin{cases} x = at + c & \text{(24.22a)} \\ y = bt + d & \text{(24.22b)} \\ z = 1. & \text{(24.22c)} \end{cases}$$

Les directions données par la droite d' sont donc

$$\frac{1}{a^2t^2 + b^2t^2 + c^2 + d^2} \begin{pmatrix} at + c \\ bt + d \\ 1 \end{pmatrix} \tag{24.23}$$

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 [3]). □

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.

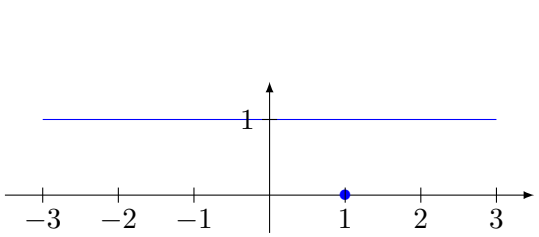
Sur la figure 24.1(a), le point à l'infini est la direction $(1, 0)$ tandis que la direction $(1, 1)$ n'a rien de spécial. À l'inverse sur la figure 24.1(b), la direction à l'infini est $(1, 1)$ tandis que la direction $(1, 0)$ est une direction usuelle.

Remarque 24.17.
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.

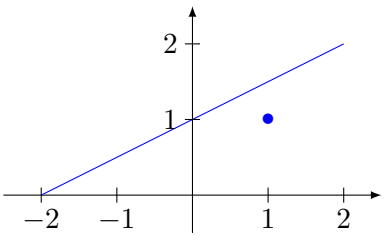
Du point de vue de la géométrie différentielle, c'est la même chose. En effet nous pouvons mettre sur la droite projective un système de deux cartes en pensant aux angles. La première sur $]-a, a[$ avec par exemple $a < \pi/4$. La seconde carte serait $]a/2, \pi[$. Dans ce cas la direction $\theta = 0$ 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).

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.



(a) Ici le point à l'infini est la direction $(1,0)$.



(b) Ici le point à l'infini est la direction $(1,1)$.

FIGURE 24.1 – 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'équivallence), en réalité les deux dessins représentent les mêmes ensembles.

24.3 Théorème de Pappus

Théorème 24.18.
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 λ_1 et λ_2 tels que

$$\begin{cases} A = \lambda_1 B & (24.24a)\\ B' = \lambda_1 A' & (24.24b) \end{cases}$$

et

$$\begin{cases} B' = \lambda_2 C' & (24.25a)\\ C = \lambda_2 B. & (24.25b) \end{cases}$$

En substituant nous trouvons

$$\begin{cases} C = \dfrac{\lambda_2}{\lambda_1} A & (24.26a)\\ A' = \dfrac{\lambda_2}{\lambda_1} C', & (24.26b) \end{cases}$$

ce qui implique que $A'C \parallel AC'$.

Si les droites d et d' sont parallèles, alors nous avons les translations

$$\begin{cases} B = A + x & (24.27a)\\ A' = B' + x & (24.27b) \end{cases}$$

et

$$\begin{cases} B = C + y & (24.28a)\\ C' = B' + y, & (24.28b) \end{cases}$$

ce qui montre que

$$\begin{cases} C = A + x - y & (24.29a)\\ A' = C' + x - y, & (24.29b) \end{cases}$$

et donc que $A'C \parallel AC'$. □

Le théorème suivant est une version projective.

Théorème 24.19.
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 $E = BC' \cap C'B$ 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'. □

24.4 Homographies

24.4.1 Homographies

Définition 24.20.
Soient E et F deux espaces vectoriels avec leurs projections naturelles

$$\pi_E \colon E\backslash\{0\} \to P(E) \tag{24.30a}$$
$$\pi_F \colon F\backslash\{0\} \to P(F). \tag{24.30b}$$

Une application $g \colon P(E) \to P(F)$ est une ***homographie*** *s'il existe un isomorphisme d'espaces vectoriels $\bar{g} \colon E \to F$ tel que le diagramme*

$$\begin{array}{ccc} E\backslash\{0\} & \xrightarrow{\bar{g}} & F\backslash\{0\} \\ {\scriptstyle \pi_E}\downarrow & & \downarrow{\scriptstyle \pi_F} \\ P(E) & \xrightarrow[g]{} & P(F) \end{array} \tag{24.31}$$

commute, c'est-à-dire s'il existe $\bar{g} \colon E \to F$ telle que

$$\pi_F\big(\bar{g}(v)\big) = g\big(\pi_E(v)\big) \tag{24.32}$$

pour tout $v \in E$.

Lemme 24.21.
Si $\bar{g} \colon E \to F$ est linéaire et si $\ker \bar{g} = \{0\}$ alors l'application g définie par

$$g\big(\pi_E(v)\big) = \pi_F\big(\bar{g}(v)\big) \tag{24.33}$$

est une homographie.

Démonstration. Nous devons simplement vérifier que l'équation (24.33) définit bien une application. Soient $v, w \in E$ tels que $\pi_E v = \pi_E w$; nous devons montrer que

$$\pi_F \bar{g} v = \pi_F \bar{g} w. \tag{24.34}$$

L'équation (24.34) sera vérifiée si et seulement s'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 (24.34) sera vérifiée si et seulement si $v = \lambda w$, ce qui signifie exactement $\pi_E(v) = \pi_E(w)$. □

La proposition suivante donne les premières propriétés des homographies.

Proposition 24.22.
Quelques propriétés des homographies.

(1) Une homographie est bijective.

(2) Si deux espaces projectifs sont homographes, alors ils ont même dimension.

(3) *L'ensemble des homographies $P(E) \to P(F)$ est un groupe (pour la composition).*

(4) *Une homographie conserve l'alignement des points.*

Démonstration. Nous considérons une homographie $g\colon P(E) \to P(F)$, et $\bar{g}$ l'isomorphisme d'espaces vectoriels correspondant.

(1) Pour l'injectivité, si $g\big([v]\big) = g\big([w]\big)$ 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 $g\big(\pi_E v\big) = \pi_F \bar{g} v$. Par conséquent l'élément $\pi_F \bar{g} v$ est bien dans l'image de g.

(2) Une homographie $P(E) \to P(F)$ n'existe que s'il existe un isomorphisme $E \to F$. Les dimensions sont donc automatiquement égales.

(3) Il suffit de vérifier que l'application

$$\begin{aligned} \varphi\colon P(E) &\to P(E) \\ \pi_F \bar{g} v &\mapsto \pi_E v \end{aligned} \tag{24.35}$$

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 affin, les points $\bar{g}a$, $\bar{g}b$ et $\bar{g}c$ sont alignés dans F. Cela implique que les projections par π_F sont alignés dans $P(F)$.

□

24.4.2 Le groupe projectif

Définition 24.23.
Le groupe des homographies de l'espace $P(E)$ est le ***groupe projectif****, noté* $\mathrm{PGL}(E)$.

Nous avons une surjection naturelle

$$\begin{aligned} \mathrm{GL}(E) &\to \mathrm{PGL}(E) \\ \bar{g} &\mapsto g \end{aligned} \tag{24.36}$$

qui s'avère être un morphisme de groupes.

Proposition 24.24.
Nous avons l'isomorphisme de groupes

$$\frac{\mathrm{GL}(E)}{\{\textit{homothéties}\}} \simeq \mathrm{PGL}(E). \tag{24.37}$$

Démonstration. Nous devons prouver que le noyau de l'application (24.36) est constitué des homothéties. Considérons un automorphisme d'espace vectoriel $f\colon E \to E$ dont l'homographie associée est l'identité, et prouvons que f est une homothétie. Nous avons le diagramme commutatif suivant :

$$\begin{array}{ccc} E\backslash\{0\} & \xrightarrow{f} & E\backslash\{0\} \\ {\scriptstyle \pi_E}\downarrow & & \downarrow{\scriptstyle \pi_E} \\ P(E) & \xrightarrow[\mathrm{Id}]{} & P(E). \end{array} \tag{24.38}$$

Pour tout vecteur $v \in E$ nous avons $\pi_E(v) = \pi_E\big(f(v)\big)$. 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. □

24.4.3 Repères projectifs

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 ne sont donc pas intrinsèques.

Définition 24.25 ([374]).
*Des éléments $\{P_i\}_{i\in I}$ sont **projectivement independents** si en choisissant $v_i \in \pi^{-1}(P_i)$ nous obtenons des vecteurs $\{v_i\}_{i\in I}$ linéairement indépendants.*

Définition 24.26.
*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 24.27
Un repère projectif de l'espace $P(\mathbb{R}^3)$ est par exemple les éléments $\{m_i\}_{i=1,\dots,3}$ donnés par

$$m_1 = \pi(e_1) \tag{24.39a}$$
$$m_2 = \pi(e_2) \tag{24.39b}$$
$$m_3 = \pi(e_3) \tag{24.39c}$$
$$m_0 = \pi(e_1 + e_2 + e_3). \tag{24.39d}$$

△

Exemple 24.28
Pour $P(\mathbb{C}^2)$, un repère projectif possible est $m_1 = [1, 0]$, $m_2 = [0, 1]$, $m_0 = [1, 1]$. △

24.29.
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 $E \to \mathbb{R}^{n+1}$ 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 $\phi\colon P(E) \to P(\mathbb{R}^{n+1})$ qui permet de donner à un point $A \in E$ les coordonnées de $\pi^{-1}\big(\phi(A)\big)$. 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 $\pi^{-1}\big(\phi(A)\big)$ 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 24.30
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 $\{e_1, e_2\}$.

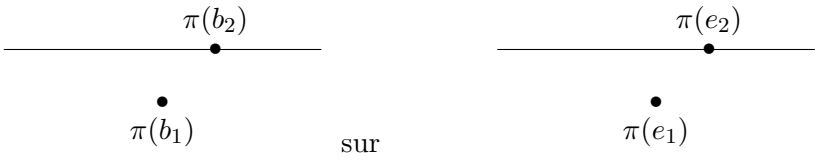
Soit une homographie $\phi\colon P(E)\to P(\mathbb{R}^2)$ telle que

$$\phi\big(\pi(b_i)\big)=\pi(e_i) \tag{24.40}$$

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 (24.40) 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 $\phi\colon P(E)\to P(\mathbb{R}^2)$, et nous définissons

$$\begin{aligned}\phi'\colon P(E)&\to P(\mathbb{R}^2)\\ \pi(xb_1+yb2)&\phi\big(\pi(xb_1+\lambda yb_2)\big)\end{aligned} \tag{24.41}$$

pour un certain $\lambda\neq1$. En ce qui concerne le relèvement, l'application $\bar{\phi}'\colon E\to\mathbb{R}^2$ donnée par

$$\bar{\phi}'(xb_1+yb_2)=\bar{\phi}(xb_1+\lambda yb_2) \tag{24.42}$$

est bien définie et vérifie

$$\pi_{\mathbb{R}^2}\circ\bar{\phi}=\phi\circ\pi_E. \tag{24.43}$$

Donc ϕ' est une homographie. De plus

$$\phi'\big(\pi(b_1)\big)=\phi\big(\pi(b_1)\big) \tag{24.44a}$$
$$\phi'\big(\pi(b_2)\big)=\phi\big(\pi(\lambda b_2)\big)=\phi\big(\pi(b_2)\big) \tag{24.44b}$$

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 $\phi\colon P(E)\to P(\mathbb{R}^{n+1})$ vérifiant $\phi\big(\pi_E(b_i)\big)=\pi_{\mathbb{R}^{n+1}}(e_i)$ pour tout $i=0,\dots,n+1$.

Lemme 24.31.
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

$$\pi(e_1+\dots+e_{n+1})=\pi(f_1+\dots+f_{n+1}) \tag{24.45}$$

alors les deux bases sont proportionnelles : il existe λ 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 λ_i sont en réalité tous égaux.

Pour cela nous avons aussi l'égalité pour $i=0$:

$$\pi(e_1+\dots+e_{n+1})=\pi(f_1+\dots+f_{n+1}), \tag{24.46}$$

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

$$\lambda_1f_1+\dots+\lambda_{n+1}f_{n+1}=\mu f_1+\dots+\mu f_{n+1}. \tag{24.47}$$

Du fait que les f_i forment une base, cette égalité impose à tous les λ_i d'être égal à μ. □

Théorème 24.32 ([375]).
Soient $P(E)$ et $P(F)$ deux espaces projectifs de dimensions n.

(1) Une homographie $P(E) \to P(F)$ 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 $g\colon P(E) \to P(F)$ telle que $g(m_i) = m'_i$ pour tout $i = 0, 1, \dots, n+1$

Démonstration. Un point à la fois.

(1) Soit une homographie $\phi\colon P(E) \to P(F)$ 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 $m'_i = \phi\big(\pi_E(e_i)\big) = \pi_F\big(\bar{\phi}(e_i)\big)$, 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

$$m'_0 = \phi\big(\pi_E(e_1 + \dots + e_{n+1})\big) = \pi_F\big(\bar{\phi}(e_1 + \dots + e_{n+1})\big) = \pi_F(f_1 + \dots + f_{n+1}), \tag{24.48}$$

ce qui termine de montrer que $\{m'_i\}_{i=0,\dots,n+1}$ est un repère projectif de $P(F)$.

(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

$$m_i = \pi_E(e_i) \tag{24.49a}$$
$$m'_i = \pi_F(f_i) \tag{24.49b}$$

pour $i = 1, \dots, n+1$ et

$$m_0 = \pi_E(e_1 + \dots + e_{n+1}) \tag{24.50a}$$
$$m'_0 = \pi_F(f_1 + \dots + f_{n+1}). \tag{24.50b}$$

Nous considérons un isomorphisme d'espace vectoriel $\bar{\phi}\colon E \to F$ tel que $\bar{\phi}(e_i) = f_i$ pour tout i, et nous voulons définir $\phi\colon P(E) \to P(F)$ par

$$\phi\big(\pi_E(v)\big) = \pi_F\big(\bar{\phi}(v)\big). \tag{24.51}$$

Cela est bien définit parce que si $\pi_E(v) = \pi_E(w)$ alors $w = \lambda v$ et

$$\pi_F\big(\bar{\phi}(\lambda v)\big) = \pi_F\big(\lambda\bar{\phi}(v)\big) = \pi_F\big(\bar{\phi}(v)\big). \tag{24.52}$$

L'application définie par (24.51) 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

$$\phi_1\colon P(E) \to P(F) \tag{24.53a}$$
$$\phi_2\colon P(E) \to P(F) \tag{24.53b}$$

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

$$\phi_1(m_i) = \pi_E\big(\bar{\phi}_1(e_i)\big) \tag{24.54a}$$
$$\phi_2(m_i) = \pi_E\big(\bar{\phi}_2(e_i)\big). \tag{24.54b}$$

Mais nous savons que $\phi_1(m_i) = \phi_2(m_i)$, donc nous savons que $\pi_E\big(\bar{\phi}_1(e_i)\big) = \pi_E\big(\bar{\phi}_2(e_i)\big)$, ce qui nous fait conclure que

$$\bar{\phi}_1(e_i) = \lambda_i \bar{\phi}_2(e_i) \tag{24.55}$$

pour certaines constantes $\lambda_i \in \mathbb{K}$. Le même raisonnement appliqué à m_0 nous donne un $\mu \in \mathbb{K}$ tel que

$$\bar{\phi}_1(e_1) + \ldots + \bar{\phi}_1(e_{n+1}) = \mu\big(\bar{\phi}_2(e_1) + \ldots + \bar{\phi}_2(e_{n+1})\big). \tag{24.56}$$

En mettant l'un dans l'autre :

$$\lambda_1\bar{\phi}_2(e_1) + \ldots + \lambda_{n+1}\bar{\phi}_2(e_{n+1}) = \mu\big(\bar{\phi}_2(e_1) + \ldots + \bar{\phi}_2(e_{n+1})\big). \tag{24.57}$$

Sachant que $\{\bar{\phi}_2(e_i)\}_{i=1,\ldots,n+1}$ est une base de F et nous souvenant de l'unicité de la décomposition d'un élément dans une base[4], nous en déduisons que tous les λ_i doivent être égaux à μ. Donc pour tout $v \in E$ nous avons $\bar{\phi}_1(v) = \lambda\bar{\phi}_2(v)$.
Cela a pour conséquence que $\phi_1 = \phi_2$.

□

24.33.
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 $f\colon P(E) \to P(E)$ et $\bar{f}\colon E \to E$, 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

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{M}(2, \mathbb{K}) \tag{24.58}$$

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

$$\bar{f}(z, 1) = (a_{11}z + a_{12}, a_{21}z + a_{22}). \tag{24.59}$$

Nous posons $\lambda = a_{21}z + a_{22}$ et nous avons

$$\bar{f}(z, 1) = \lambda\left(\frac{a_{11}z + a_{12}}{\lambda}, 1\right). \tag{24.60}$$

Il y a plusieurs possibilités suivant les valeurs de λ 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.
(2) Si $\lambda \neq$, 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

$$\begin{cases} \varphi_f(z) = \dfrac{a_{11}z + a_{12}}{a_{21}z + a_{22}} & \text{(24.61a)} \\ \varphi_f(\infty) = \dfrac{a_{11}}{a_{21}} & \text{(24.61b)} \end{cases}$$

alors nous avons

$$\bar{f}(z, 1) = \big(\varphi_f(z), 1\big). \tag{24.62}$$

Si nous prenons la convention que $\frac{1}{0} = \infty$ et que $(\infty, 0)$ est le point à l'infini, alors cette application φ_f donne bien toutes les valeurs de f, y compris les cas à l'infini.

4. Proposition 4.6.

Lemme 24.34 ([376]).
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 [5].

Il existe donc $\alpha, \beta \in \mathbb{K}$ tels que $c = \alpha a + \beta b$. De plus α et β 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 $d = P(E)$. □

Corollaire 24.35 ([377]).
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 $P(E) \to P(F)$ portant A_i sur B_i pour tout $i = 1, 2, 3$.

Démonstration. Il s'agit de mettre en conjonction le lemme 24.34 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 24.32(2) qui dit que l'une va sur l'autre par une unique homographie. □

24.4.4 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 24.8).

24.36.
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

$$\begin{aligned} \varphi_0 \colon P(\mathbb{K}^2) &\to \mathbb{K} \cup \{\infty\} \\ [k_1, k_2] &\mapsto \begin{cases} \frac{k_1}{k_2} & \text{si } k_2 \neq 0 \\ \infty & \text{si } k_2 = 0. \end{cases} \end{aligned} \tag{24.63}$$

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.

Corollaire 24.37.
Soit une bijection $\varphi \colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$. Les points $\{\varphi^{-1}(\infty), \varphi^{-1}(0), \varphi^{-1}(1)\}$ forment un repère projectif de $P(\mathbb{K}^2)$.

Démonstration. Il s'agit seulement d'une application du lemme 24.34.

Juste pour l'amusement, nous allons le prouver explicitement pour la bijection $\varphi = \varphi_0$ donnée en (24.63). Un repère projectif est la définition 24.26. Nous avons

$$\varphi_0^{-1}(\infty) = [1, 0] = \pi_{K^2}\big((1, 0)\big) \tag{24.64a}$$
$$\varphi_0^{-1}(0) = [0, 1] = \pi_{K^2}\big((0, 1)\big) \tag{24.64b}$$
$$\varphi_0^{-1}(1) = [1, 1] = \pi_{K^2}\big((1, 1)\big) \tag{24.64c}$$

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. □

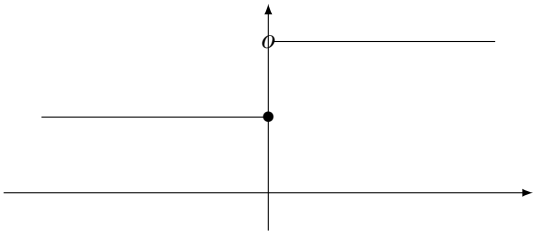
D'autre part, comme il est plus agréable de travailler avec $\hat{\mathbb{K}}$ qu'avec $P(\mathbb{K}^2)$ nous avons envie de voir $\hat{\mathbb{K}}$ comme un espace projectif (qu'il n'est pas). Il y a cependant nombre d'autres identifications

5. Il convient de citer ici la proposition 4.16.

possibles. En voici un autre :

$$\begin{aligned}\varphi_1\colon P(\mathbb{K}^2) &\to \hat{\mathbb{K}}\\ [k_1,k_2] &\mapsto \begin{cases}\frac{k_2}{k_1} & \text{si } k_1\neq 0\\ \infty & \text{si } k_1=0.\end{cases}\end{aligned} \tag{24.65}$$

Vous en voulez une plus compliquée ? En voici une pour $\mathbb{K}=\mathbb{R}$, basée sur le dessin suivant :



La bijection associée est :

$$\begin{aligned}\varphi_d\colon P(\mathbb{R}^2) &\to \hat{\mathbb{R}}\\ [k_1,k_2] &\mapsto \begin{cases}\frac{k_1}{k_2} & \text{si } k_1k_2\leqslant 0\\ 2\frac{k_1}{k_2} & \text{si } k_1k_2<0\\ \infty & \text{si } k_2=0.\end{cases}\end{aligned} \tag{24.66}$$

Pour mettre un peu d'ordre dans toutes ces identifications possibles, nous introduisons une classe.

Définition 24.38.
Pour une bijection $\varphi\colon P(\mathbb{K}^2)\to\hat{\mathbb{K}}$ nous définissons

$$A(\varphi)=\{\varphi_a\colon P(\mathbb{K}^2)\to\hat{\mathbb{K}} \text{ tel que } \varphi_a^{-1}\circ\varphi \text{ soit une homographie }\} \tag{24.67}$$

Les classes sont assez larges parce que pour toute homographie $\phi\colon P(\mathbb{K}^2)\to P(\mathbb{K}^2)$, nous avons $\varphi\circ\phi^{-1}\in A(\varphi)$. Mieux, nous avons le lemme suivant.

Lemme 24.39.
L'application

$$\begin{aligned}\psi\colon A(\varphi) &\to \mathrm{PGL}(\mathbb{K}^2)\\ \varphi_a &\mapsto \varphi_a^{-1}\circ\varphi\end{aligned} \tag{24.68}$$

est une bijection.

Démonstration. Pour rappel, $\mathrm{PGL}(E)$ est le groupe des homographies de E, voir la définition 24.23.

Surjectif Si $\phi\in\mathrm{PGL}(\mathbb{K}^2)$ nous avons $\phi=\psi(\varphi\circ\phi^{-1})$.

Injectif Si $\psi(\varphi_a)=\psi(\varphi_b)$ alors

$$\varphi_a^{-1}\circ\varphi=\varphi_b^{-1}\circ\varphi, \tag{24.69}$$

d'où nous déduisons $\varphi_a^{-1}=\varphi_b^{-1}$ parce que φ est une bijection.

□

24.4.5 Birapport

Proposition-définition 24.40.
*Soit une droite projective $d=P(E)$ et trois points distincts A, B et C sur cette droite. Soit une bijection $\varphi\colon P(\mathbb{K}^2)\to\hat{\mathbb{K}}$. 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*

$$[A,B,C,X]_\varphi=(\varphi\circ\phi)(X) \tag{24.70}$$

où $\phi\colon d \to P(\mathbb{K}^2)$ *est l'unique homographie telle que*

$$\phi(A) = \varphi^{-1}(\infty) \tag{24.71a}$$
$$\phi(B) = \varphi^{-1}(0) \tag{24.71b}$$
$$\phi(C) = \varphi^{-1}(1). \tag{24.71c}$$

Démonstration. Nous devons prouver qu'il existe effectivement une unique homographie vérifiant les conditions (24.71).

A, B, C **est un repère projectif de** $P(E)$ Voir le lemme 24.34.

$\varphi^{-1}(\infty), \varphi^{-1}(0), \varphi^{-1}(1)$ **est un repère projectif de** $P(\mathbb{K}^2)$ Lemme 24.34 ou corollaire 24.37 au choix.

Conclusion Le théorème 24.32(2) nous donne existence et unicité d'une homographie $P(E) \to P(\mathbb{K}^2)$ envoyant le premier repère sur le second.

□

Remarque 24.41.
La majorité des sources ne parlent pas de la dependence du birapport en le choix de φ parce que tout le monde ne semble ne considérer que $\varphi = \varphi_0$ définie en (24.63). Il est cependant naturel de se demander si la définition dépend effectivement du choix de φ. La réponse est oui : ça dépend du choix.

Exemple 24.42(Une autre identification qui ne va pas bien)
Nous montrons que l'identification $\varphi_d\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$ donnée en (24.66) ne donne pas lieu au même birapport que celui de φ_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 $[.,.,.,.]_{\varphi_0}$ est $\varphi_0 = \mathrm{Id}$. Par conséquent,

$$\big[A, B, C, [k_1, k_2]\big]_{\varphi_0} = (\varphi_0 \circ \phi_0)[k_1, k_2] = \varphi_0[k_1, k_2] = \frac{k_1}{k_2}. \tag{24.72}$$

En ce qui concerne le birapport définit par φ_d nous avons

$$\varphi_d^{-1}(\infty) = [1, 0] \tag{24.73a}$$
$$\varphi_d^{-1}(0) = [0, 1] \tag{24.73b}$$
$$\varphi_d^{-1}(1) = [1, 2], \tag{24.73c}$$

de telle sorte que nous cherchons une homographie $\phi_d\colon P(\mathbb{K}^2) \to P(\mathbb{K}^2)$ telle que

$$\phi_d[1, 0] = [1, 0] \tag{24.74a}$$
$$\phi_d[0, 1] = [0, 1] \tag{24.74b}$$
$$\phi_d[1, 1] = [1, 2] \tag{24.74c}$$

L'homographie $\phi_d[k_1, k_2] = [k_1, 2k_2]$ convient et nous avons

$$\big[A, B, C, D, [k_1, k_2]\big]_{\varphi_d} = (\varphi_d \circ \phi_d)[k_1, k_2] = \varphi_d[k_1, 2k_2] = \frac{k_1}{2k_2} \tag{24.75}$$

dès que $k_1 k_2 < 0$. Nous avons donc bien trouvé

$$[A, B, C, X]_{\varphi_0} \neq [A, B, C, X]_{\varphi_d}. \tag{24.76}$$

△

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 $\hat{\mathbb{K}}$.

Proposition 24.43.
Soit une bijection $\varphi\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$. *Si* $\varphi_a \in A(\varphi)$[6] *alors les birapports construits sur* φ *et* φ_a

6. $A(\varphi)$ définie en 24.38.

coïncident.

Démonstration. Soient trois points distincts $A, B, C \in P(E)$, et $X \in P(E)$. Nous avons

$$[A, B, C, X]_\varphi = (\varphi \circ \phi)(X) \tag{24.77}$$

où $\phi\colon P(E) \to P(\mathbb{K}^2)$ est l'unique homographie telle que

$$\phi(A) = \varphi^{-1}(\infty) \tag{24.78a}$$
$$\phi(B) = \varphi^{-1}(0) \tag{24.78b}$$
$$\phi(C) = \varphi^{-1}(1). \tag{24.78c}$$

Et

$$[A, B, C, X]_{\varphi_a} = (\varphi_a \circ \phi_a)(X) \tag{24.79}$$

où $\phi_a\colon P(E) \to P(\mathbb{K}^2)$ est l'unique homographie telle que

$$\phi_a(A) = \varphi_a^{-1}(\infty) \tag{24.80a}$$
$$\phi_a(B) = \varphi_a^{-1}(0) \tag{24.80b}$$
$$\phi_a(C) = \varphi_a^{-1}(1). \tag{24.80c}$$

Il est facile de voir que $\phi_a = \varphi_a^{-1} \circ \varphi\phi$. 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,

$$(\varphi_a^{-1} \circ \varphi \circ \phi)(A) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(\infty) = \varphi_a(\infty) \tag{24.81a}$$
$$(\varphi_a^{-1} \circ \varphi \circ \phi)(B) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(0) = \varphi_a(0) \tag{24.81b}$$
$$(\varphi_a^{-1} \circ \varphi \circ \phi)(C) = (\varphi_a^{-1} \circ \varphi)\phi^{-1}(1) = \varphi_a(1). \tag{24.81c}$$

Au final nous avons :

$$[A, B, C, X]_{\varphi_a} = (\varphi_a \circ \varphi_a^{-1} \circ \varphi \circ \phi)(X) = (\varphi \circ \phi)(X) = [A, B, C, X]_\varphi. \tag{24.82}$$

□

Remarque 24.44.
Tout le monde semble ne considérer que l'identification usuelle $\varphi_0\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$ 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 considéra que φ_0.

Et c'est ce que nous allons faire : sauf avis contraire, nous utiliserons le birapport associé à l'identification φ_0.

Lemme 24.45 ([375])**.**
Nous avons

$$[A, B, C, X]_\varphi = \begin{cases} \infty & \text{si et seulement si } X = A \\ 0 & \text{si et seulement si } X = B \\ 1 & \text{si et seulement si } X = C. \end{cases} \tag{24.83}$$

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 ϕ et φ sont des bijections.

Le même raisonnement tient pour les deux autres. □

Proposition 24.46.
Autres petites propriétés faciles ... Soit une droite projective $d = P(E)$ et trois points distincts $A, B, C \in d$.

(1) Les points A, B, C et X sont distincts si et seulement si $[A, B, C, X] \in \mathbb{K}\backslash\{0, 1\}$.

(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 définit.

(2) Les points A, B et C sont distincts par hypothèse. Vu le lemme 24.45, pour que X soit distincts de A, B et C il faut et il suffit que le birapport ne soit ni ∞ ni 1 ni 0. Donc $\mathbb{K}\backslash\{0, 1\}$.

(2) Nous avons $[A, B, C, X] = \phi(X)$ où $\phi\colon P(E) \to P(\mathbb{K}^2)$ est une homographie et donc une bijection par la proposition 24.22(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 (24.63). □

Proposition 24.47 ([375])**.**
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

(1) Il existe une homographie $\phi\colon d \to d'$ telle que $\phi(A_i) = A'_i$ pour $i = 1, 2, 3, 4$,

(2) égalité des birapports :

$$[A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4]. \tag{24.84}$$

Dans ce cas, l'homographie est unique.

Démonstration. Nous divisons la preuve en trois parties évidentes.

(1) implique (2) Nous avons une homographie $\mu'\colon d' \to \hat{\mathbb{K}}$ telle que $\mu'(A'_1) = \infty$, $\mu'(A'_2) = 0$ et $\mu'(A'_3) = 1$. En composant [7] avec l'homographie $\phi\colon d \to d'$ de l'hypothèse nous avons une homographie $\mu' \circ \phi\colon d \to \hat{\mathbb{K}}$ qui vérifie

$$(\mu' \circ \phi)(A_1) = \mu'(A'_1) = \infty \tag{24.85a}$$
$$(\mu' \circ \phi)(A_2) = \mu'(A'_2) = 0 \tag{24.85b}$$
$$(\mu' \circ \phi)(A_3) = \mu'(A'_3) = 1, \tag{24.85c}$$

ce qui signifie que $\mu' \circ \phi$ est l'homographie qui définie le birapport sur d. Par conséquent

$$[A_1, A_2, A_3, A_4] = (\mu' \circ \phi)(A_4) = \mu'(A'_4) = [A'_1, A'_2, A'_3, A'_4]. \tag{24.86}$$

(2) implique (1) La partie $\{A_1, A_2, A_3\}$ est un repère projectif de d par le lemme 24.34. Idem pour $\{A'_1, A'_2, A'_3\}$. Nous considérons les homographies $\mu\colon d \to \hat{\mathbb{K}}$ et $\mu'\colon d' \to \hat{\mathbb{K}}$ définissant les birapports par rapport à ces repères. Ces homographies vérifient, en utilisant l'hypothèse :

$$\mu(A_4) = [A_1, A_2, A_3, A_4] = [A'_1, A'_2, A'_3, A'_4] = \mu'(A'_4). \tag{24.87}$$

Et de plus

$$\mu(A_1) = [A_1, A_2, A_3, A_1] = \infty = \mu'(A'_1) \tag{24.88a}$$
$$\mu(A_2) = [A_1, A_2, A_3, A_2] = 0 = \mu'(A'_2) \tag{24.88b}$$
$$\mu(A_3) = [A_1, A_2, A_3, A_3] = 1 = \mu'(A'_3). \tag{24.88c}$$

7. Proposition 24.22(3), la composition est encore une homographie.

Autrement dit : $\mu(A_i) = \mu'(A'_i)$ pour tout $i = 1, 2, 3, 4$. Nous nous inspirons de ce diagramme :

$$\begin{array}{ccc} d & \overset{\phi}{\dashrightarrow} & d' \\ & {\scriptstyle\mu}\searrow \quad \swarrow{\scriptstyle\mu'} & \\ & \mathbb{K} \cup \{\infty\} & \end{array} \tag{24.89}$$

La composée $\phi = \mu'^{-1} \circ \mu$ vérifie

$$(\mu'^{-1} \circ \mu)(A_i) = A'_i \tag{24.90}$$

pour tout i, et est une homographie.

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 24.48 ([375]).
Une bijection entre deux droites projectives est une homographie si et seulement si elle conserve le birapport.

Démonstration. Chacun des deux sens séparément.

⇒ Soit une homographie $\phi\colon d \to d'$ 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

$$[A, B, C, X] = [\phi(A), \phi(B), \phi(C), \phi(X)]. \tag{24.91}$$

Nous nommons $\mu\colon d \to \hat{\mathbb{K}}$ l'homographie qui donne le birapport sur d par rapport à A, B et C, et $\mu'\colon d' \to \hat{\mathbb{K}}$ celle qui donne le birapport sur d' par rapport à $\phi(A)$, $\phi(B)$, $\phi(C)$. Voici un diagramme de la situation :

$$\begin{array}{ccc} d & \overset{\phi}{\dashrightarrow} & d' \\ & {\scriptstyle\mu}\searrow \quad \swarrow{\scriptstyle\mu'} & \\ & \mathbb{K} \cup \{\infty\} & \end{array} \tag{24.92}$$

Nous prouvons maintenant que $\mu' = \mu \circ \phi^{-1}$. En effet :

$$(\mu \circ \phi^{-1})\big(\phi(A)\big) = \mu(A) = \infty \tag{24.93a}$$
$$(\mu \circ \phi^{-1})\big(\phi(B)\big) = \mu(B) = 0 \tag{24.93b}$$
$$(\mu \circ \phi^{-1})\big(\phi(C)\big) = \mu(C) = 1. \tag{24.93c}$$

Par conséquent le birapport à droite dans (24.91) peut se calculer à l'aide de $\mu \circ \phi^{-1}$:

$$[\phi(A), \phi(B), \phi(C), \phi(X)] = \mu'\big(\phi(X)\big) = (\mu \circ \phi^{-1})\big(\phi(X)\big) = \mu(X) = [A, B, C, X]. \tag{24.94}$$

La première implication est prouvée.

⇐ Soit une bijection $f\colon d \to d'$ 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 24.34 et le théorème 24.32(2), il existe une unique homographie $\phi\colon d \to d'$ telle que $\phi(A) = f(a)$, $\phi(B) = f(B)$ et $\phi(C) = f(C)$. Pour tout $X \in d$ nous avons

$$[f(A), f(B), f(C), f(X)] = [A, B, C, X] \tag{24.95a}$$
$$= [\phi(A), \phi(B), \phi(C), \phi(X)] \tag{24.95b}$$
$$= [f(A), f(B), f(C), \phi(X)]. \tag{24.95c}$$

Justifications :

— (24.95a) parce que f conserve le birapport par hypothèse.

— (24.95b) parce que ϕ conserve le birapport étant une homographie (c'est le premier sens du présent théorème)

Nous nommons $\mu' : d' \to \hat{\mathbb{K}}$ l'homographie donnant le birapport par rapport aux points $f(A)$, $f(B)$, $f(C)$. Alors le résultat (24.95) se lit

$$\mu'\big(f(X)\big) = \mu'\big(\phi(X)\big). \tag{24.96}$$

Mais comme μ' est une bijection (proposition 24.22(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 24.49.
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

$$d = \pi_E(\lambda x + \mu y) \tag{24.97}$$

si et seulement si

$$[a, b, c, d] = \pi_{\mathbb{K}^2}(\lambda, \mu). \tag{24.98}$$

Démonstration. Étant donné que a et b sont distincts, les vecteurs x et y forment une base de E. Soit $f : E \to \mathbb{K}^2$ 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 $g : P(E) \to P(\mathbb{K}^2)$, l'homographie associée à f. Par définition $f\big(\pi_E z\big) = \pi_{\mathbb{K}^2}\big(f(z)\big)$. Par f nous avons

$$a \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad b \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{24.99}$$

Donc par g nous avons

$$a \mapsto \infty \quad b \mapsto 0. \tag{24.100}$$

Nous avons aussi $f(\lambda x + \mu y) = (\lambda, \mu)$ et

$$\begin{aligned} g(c) &= g\big(\pi_E(x + y)\big) && \text{(24.101a)} \\ &= \pi_F f(x + y) && \text{(24.101b)} \\ &= \pi_F(f(x) + f(y)) && \text{(24.101c)} \\ &= \pi_F \begin{pmatrix} 1 \\ 1 \end{pmatrix} && \text{(24.101d)} \\ &= 1. && \text{(24.101e)} \end{aligned}$$

La dernière inégalité 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

$$g(d) = \pi_{\mathbb{K}^2} f(\lambda x + \mu y) = \pi_{\mathbb{K}^2}(\lambda, \mu). \tag{24.102}$$

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

$$g\pi_E v = \pi_{\mathbb{K}^2} f(v), \tag{24.103}$$

ce qui implique $f(v) = \alpha(\lambda, \mu)$ pour un certain $\alpha \in \mathbb{K}$. Par conséquent $v = \alpha(\lambda x + \mu y)$ et $d = \pi_E(\lambda x + \mu y)$. □

24.5 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 $x_i=\lambda x_i$.

Définition 24.50.
La classe d'équivalence de $(x_0,\dots,x_n)$ est la ***coordonnées homogène*** *de M. Nous la notons $(x_0:\dots:x_n)$.*

24.5.1 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 ω est un hyperplan. Le noyau de la forme linéaire $\lambda\omega$ étant le même hyperplan, l'hyperplan est donné par toute la classe de ω dans $P(E^*)$. Nous avons donc une bijection

$$P(E^*)\leftrightarrow\{\text{hyperplans vectoriels de } E\}. \tag{24.104}$$

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

$$H_m\{(X:Y:T)\text{ tel que }m(X,Y,T)=0\} \tag{24.105}$$

dans $P(E)$. Si les coordonnées homogènes de m étaient $(u:v:w)$ alors l'équation de la droite H_m est

$$uX+vY+wT=0. \tag{24.106}$$

En effet si $\omega\in E^*$ est un représentant de m alors $\omega=\lambda(ue_1^*+ve_2^*+we_3^*)$ et l'équation (24.106) est indépendante de λ 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

$$\lambda m_1(X,Y,T)+\mu m_2(X,Y,T)=0 \tag{24.107}$$

qui sont encore des droites dans $P(E)$. Toutes ces droites passent par le point d'intersection des droits associées à m_1 et m_2. Nous avons donc

$$\bigcap_{\lambda,\mu}H_{\lambda m_1+\mu m_2}=H_{m_1}\cap H_{m_2}. \tag{24.108}$$

Lemme 24.51.
L'application

$$\begin{aligned}P(E^*)&\to\{\textit{droites dans } P(E)\}\\ m&\mapsto H_m\end{aligned} \tag{24.109}$$

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 ω_1 et ω_2 associées dans E^* ne sont pas multiples l'une de l'autre. Donc les équations

$$a_1X+b_1Y+z_1T=0 \tag{24.110}$$

et

$$a_2X + b_2Y + z_2T = 0 \tag{24.111}$$

n'ont pas de solutions communes et décrivent donc des droites distinctes. □

Lemme 24.52.
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

$$m_3 = m_1 + \mu(m_2 - m_1). \tag{24.112}$$

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 (24.112) 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

$$\begin{cases} a_1X + b_1Y + c_1T = 0 & \text{(24.113a)} \\ a_2X + b_2Y + c_2T = 0 & \text{(24.113b)} \\ a_3X + b_3Y + c_3T = 0. & \text{(24.113c)} \end{cases}$$

Afin que cela ait une solution non triviale nous devons avoir

$$\det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \neq 0, \tag{24.114}$$

c'est-à-dire que les points (a_i, b_i, c_i) soient alignés. □

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 $P(E)$.

24.5.2 Polynômes

Soit l'espace projectif de dimension n avec ses coordonnées homogènes $(X_0 : \ldots : X_n)$. 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 $(0, \ldots, 0, 1)$. Considérons un polynôme homogène P sur le corps $\mathbb{K}$. L'équation

$$P(X_0, \ldots, X_n) = 0 \tag{24.115}$$

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 $P(X_0, \ldots, X_n) = 0$. 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

$$Q(x_0, \ldots, x_{n-1}) = 0 \tag{24.116}$$

où Q est le polynôme donné par $Q(X_0, \ldots, x_{n-1}) = P(x_0, \ldots, x_{n-1}, 1)$. Les points de A ayant un représentant sur la droite à l'infini s'obtiennent par l'équation

$$R(x_0, \ldots, x_{n-1}) = 0 \tag{24.117}$$

où R est le polynôme donné par $R(x_0,\dots,x_{n-1}) = P(x_0,\dots,x_{n-1},0)$.

Exemple 24.53
Nous considérons la conique projective

$$X^2 - XT - Y^2 - T^2 = 0. \tag{24.118}$$

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 (24.118) :

$$x^2 - x - y^2 - 1 = 0. \tag{24.119}$$

L'autre est obtenue en posant $T = 0$:

$$x^2 - y^2 = 0. \tag{24.120}$$

La partie à l'infini est donc composée de deux points : $(1 : 1 : 0)$ et $(1 : -1 : 0)$.

Le graphique de l'équation (24.119) est donné à la figure 24.2. Nous y voyons que les asymptotes sont effectivement données par les directions $(1, 1)$ et $(1, -1)$ dans le plan.

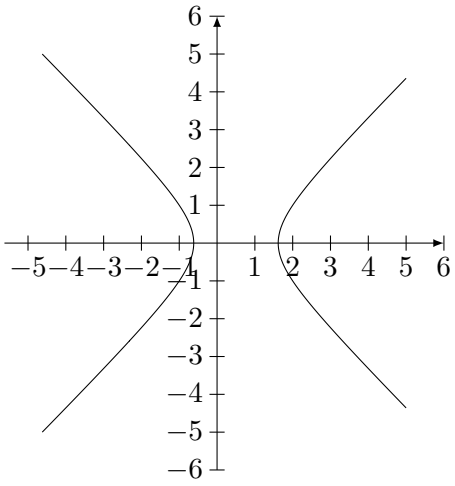


FIGURE 24.2 – Le graphique de $x^2 - x - y^2 - 1 = 0$.

△

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 24.54
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. △

Exemple 24.55
Prenons la conique

$$x^2 + xy + y^3 - 2 = 0. \tag{24.121}$$

D'abord nous homogénéisons cette équation pour la voir dans $\mathbb{R}^3$:

$$x^2z + xyz + y^3 - 2z^3 = 0. \tag{24.122}$$

Les points à l'infini sont ceux qui correspondent à $z = 0$, c'est-à-dire la droite donnée en coordonnées homogènes par $(1 : 0 : 0)$. △

24.6 La sphère de Riemann $P_1(\mathbb{C})$

Définition 24.56.
*La **sphère de Riemann** est l'espace projectif modelé sur $\mathbb{C}^2$: en vertu des notations données à la page 1507, c'est*

$$P_1(\mathbb{C}) = P(\mathbb{C}^2). \tag{24.123}$$

L'ensemble $P_1(\mathbb{C})$ est le quotient $\mathbb{C}^2\backslash\{(0,0)\}/\sim$ où $\sim$ est la relation d'équivalence de $\mathbb{C}$-colinéarité dans $\mathbb{C}^2$.

Lemme 24.57.
L'application

$$\begin{aligned}\varphi_0\colon P_1(\mathbb{C}) &\to \mathbb{C}\cup\{\infty\}\\ [z_1,z_2] &\mapsto \begin{cases}\frac{z_1}{z_2} & si\ z_2\neq 0\\ \infty & si\ z_2=0\end{cases}\end{aligned} \tag{24.124}$$

est une bijection qui respecte la conjugaisons complexe : $\varphi_0\big([z_1,z_2]^*\big)=\varphi_0\big([z_1,z_2]\big)^*$.

Démonstration. Notons d'abord que la définition a un sens parce que si un représentant que $[z_1,z_2]$ 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.

Injectif Supposons $\varphi_0\big([z_1,z_2]\big)=\varphi_0\big([t_1,t_2]\big)$.

Si les deux membres sont égaux à ∞ 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

$$z_2=\frac{t_2}{t_1}z_1, \tag{24.125}$$

et donc

$$\frac{t_1}{z_1}(z_1,z_2)=(t_1,t_2), \tag{24.126}$$

qui montre qu'au niveau des classes, $[z_1,z_2]=[t_1,t_2]$.

Surjectif Nous avons

$$\infty=\varphi_0\big([1,0]\big) \tag{24.127}$$

et si $z\neq\infty$ nous avons $z=\varphi_0\big([z,1]\big)$.

□

24.6.1 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 24.58.
La définition 24.2 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.

24.6.1.1 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[378]

$$x = \frac{z + \bar{z}}{2}, \quad y = \frac{z - \bar{z}}{2i}, \tag{24.128}$$

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

$$\bar{\omega}z + \omega z = k. \tag{24.129}$$

Définition 24.59.
*Une **droite** est une partie de $\hat{\mathbb{C}}$ de la forme*

$$d(\omega, k) = \{z \in \mathbb{C} \text{ tel que } \bar{\omega}z + \omega\bar{z} = k\} \cup \{\infty\} \tag{24.130}$$

avec $\omega \in \mathbb{C}^*$ et $k \in \mathbb{R}$.

Dans $\hat{\mathbb{C}}$, toutes les droites contiennent le point ∞.

Problèmes et choses à faire

La proposition 23.11 montre que toute inversion transforme un cercle-droite en un cercle-droite, nonobstant d'accepter de prolonger toute droite par ∞.

Est-ce que l'on peut dire que toutes les droites contiennent le point ∞ ?
En donnant ∞ à toutes les droites et à aucun cercle, la proposition 23.11 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.

24.6.1.2 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 :

$$|z - \omega| = r \Leftrightarrow |z - \omega|^2 = r^2 \Leftrightarrow (z - \omega)(\bar{z} - \bar{\omega}) = r^2 \Leftrightarrow z\bar{r} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2. \tag{24.131}$$

Donc un cercle de centre $\omega \in \mathbb{C}$ et de rayon $r \in \mathbb{R}$ a pour équation

$$z\bar{z} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2. \tag{24.132}$$

Définition 24.60.
*Un **cercle** dans $\hat{\mathbb{C}}$ est une partie de la forme*

$$C(\omega, r) = \{z \in \mathbb{C} \text{ tel que } z\bar{z} - \bar{\omega}z - \omega\bar{z} = r^2 - |\omega|^2\} \tag{24.133}$$

avec $\omega \in \mathbb{C}$ et $r \in \mathbb{R}$.

Dans la sphère de Riemann, aucun cercle ne contient le point ∞.

Exemple 24.61
Trouvons le centre et le rayon du cercle d'équation

$$\bar{\omega}z + \omega\bar{z} = kz\bar{z} \tag{24.134}$$

avec $k \neq 0$. En divisant par k et en posant $\sigma = \omega/k$ nous avons :

$$z\bar{z} - \bar{\sigma}z - \sigma\bar{z} = 0. \tag{24.135}$$

Cela est un cercle de centre σ et de rayon $|\sigma|$. En effet si $z \in \mathbb{C}$ vérifie cette équation,

$$|z - \sigma|^2 = (z - \sigma)(\bar{z} - \bar{\sigma}) = \underbrace{z\bar{z} - \bar{\sigma}z - \sigma\bar{z}}_{=0} + |\sigma|^2 = |\sigma|^2, \tag{24.136}$$

c'est-à-dire que tous les points de $\mathbb{C}$ qui vérifient l'équation donnée sont à la distance $|\sigma|$ de σ. En particulier $z = 0$ est sur le cercle. △

24.6.1.3 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 24.62 ([378]).
*Un **cercle-droite** est l'ensemble des points $z \in \mathbb{C}$ tels que*

$$az\bar{z} - \bar{\omega}z - \omega\bar{z} = k \tag{24.137}$$

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 (24.129). Si $a \neq 0$ alors nous pouvons diviser par a, poser $\sigma = \omega/a$ et $l = k/a$ pour obtenir

$$z\bar{z} - \bar{\sigma} - \sigma\bar{z} = l, \tag{24.138}$$

qui est l'équation (24.132) d'un cercle ... ou pas tout à fait. En effet, (24.138) serait l'équation du cercle de centre σ et de rayon r donné par $l = r^2 - |\sigma|^2$, c'est) dire

$$r^2 = l + |\sigma|^2, \tag{24.139}$$

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. □

Remarque 24.63.
Lorsque nous parlons de cercle-droite, nous parlons de partie de $\mathbb{C}$ et non de $\hat{\mathbb{C}}$ parce que l'équation (24.137) a du mal à traiter le cas $z = \infty$. À cause du fait que nous avons décider de donner le point ∞ 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.

Exemple 24.64([378])
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 :

$$z\bar{r} + irz - ir\bar{z} = 0. \tag{24.140}$$

Vu que $ir \neq 0$ nous pouvons diviser et obtenir

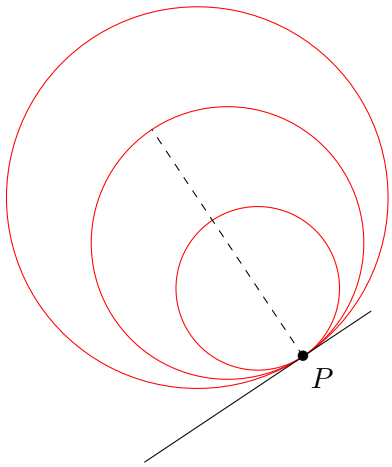
$$\frac{z\bar{z}}{ir} + z - \bar{z} = 0. \tag{24.141}$$

En faisant tendre r vers ∞ 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. △

24.65.
Notons que l'exemple 24.64 est générique : prenez une droite ℓ, un point P sur ℓ, et considérez un cercle dont le centre est situé sur la perpendiculaire à ℓ passant par P, et dont le rayon est tel que le cercle passe par P. En prenant $|\omega - P| \to \infty$, l'équation du cercle devient celle de la droite ℓ.

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 ∞.



24.6.1.4 Rotation-homothétie

Définition 24.66.
*Une **rotation-homothétie** est une application $\hat{\mathbb{C}} \to \hat{\mathbb{C}}$ de la forme $z \mapsto \lambda z$ avec $\lambda \in \mathbb{C}$.*

Le nom provient du fait que si λ 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.

24.6.1.5 Application linéaire

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

$$\begin{aligned} f\colon \hat{\mathbb{C}} &\to \hat{\mathbb{C}} \\ z &\mapsto \alpha z + \beta, \end{aligned} \tag{24.142}$$

nous entendons implicitement que $f(\infty) = \infty$.

24.6.1.6 Inversion

L'inversion d'un cercle de $\mathbb{R}^2$ est définie par la proposition 23.5. De nombreuses propriétés y sont décrites, y compris son écriture complexe dans la proposition 23.10. 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.

Définition 24.67 ([378]).
*Soient $\omega \in \mathbb{C}$ et $R \in \mathbb{R}^*$. L'**inversion** de centre ω et de **puissance** R^2 est l'application*

$$\begin{aligned} i\colon \hat{\mathbb{C}} &\to \hat{\mathbb{C}} \\ z &\mapsto \begin{cases} \dfrac{R^2}{\bar{z} - \bar{\omega}} + \omega & si\ z \in \mathbb{C}\backslash\{\omega\} \\ \infty & si\ z = \omega \\ \omega & si\ z = \infty. \end{cases} \end{aligned} \tag{24.143}$$

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 24.68
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 $\infty \mapsto 0$. △

24.6.2 Homographies

La notion d'homographie est la définition 24.20. Pour une homographie $\phi\colon P_1(\mathbb{C}) \to P_1(\mathbb{C})$ nous avons un isomorphisme d'espace vectoriel $\bar{\phi}\colon \mathbb{C}^2 \to \mathbb{C}^2$. Vue la bijection (24.124), 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

$$\begin{aligned} &\tilde{\phi}\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}} \\ &\tilde{\phi} = \varphi_0 \circ \phi \circ \varphi_0^{-1}. \end{aligned} \tag{24.144}$$

Proposition 24.69.
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 $\tilde{\phi}\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ est de la forme

$$\tilde{\phi}(z) = \frac{az + b}{cz + d} \tag{24.145}$$

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.

$\Rightarrow$ La condition $\pi \circ \bar{\phi} = \phi \circ \pi$ (de (24.32)) nous dit que

$$\big[\bar{\phi}(z_1, z_2)\big] = \phi\big([z_1, z_2]\big), \tag{24.146}$$

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

$$\bar{\phi}(z_1, z_2) = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix}. \tag{24.147}$$

Soit $z \in \mathbb{C}$. Alors nous avons

$$\tilde{\phi}(z) = (\varphi_0 \circ \phi)[z, 1] = \varphi_0\Big(\big[\bar{\phi}(z, 1)\big]\Big) = \varphi_0\big([az + b, cz + d]\big) = \frac{az + b}{cz + d}. \tag{24.148}$$

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,

$$\tilde{\phi}(-d/c) = \frac{-\frac{ad}{c} + b}{0}. \tag{24.149}$$

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,

$$\tilde{\phi}(-d/c) = \infty \tag{24.150}$$

automatiquement, et cela est encodé dans la formule (24.148).

Il nous reste à déterminer $\tilde{\phi}(\infty)$. Nous avons :

$$\tilde{\phi}(\infty) = (\varphi_0 \circ \phi)\big([1, 0]\big) = \varphi_0\big[\bar{\phi}(1, 1)\big] = \varphi_0\big([a, c]\big) = \begin{cases} a/c & \text{si } c \neq 0 \\ \infty & \text{si } c = 0 \end{cases} \tag{24.151}$$

où la distinction entre les deux cas n'est pas fondamentale parce que si $c = 0$ alors $a \neq 0$ et $a/c = \infty$.

$\Leftarrow$ En notant $Z = [z_1, z_2]$ (pour $z_1, z_2 \in \mathbb{C}$) nous avons

$$\phi\big([z_1, z_2]\big) = \varphi^{-1}\left(\frac{a\varphi_0(Z) + b}{c\varphi_0(Z) + d}\right) = \big[a\varphi_0(Z) + b, c\varphi_0(Z) + d\big]. \tag{24.152}$$

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 $\bar{\phi}\colon \mathbb{C}^2 \to \mathbb{C}^2$ dont nous avons besoin pour

la définition 24.20. D'abord le fait que $ad-cb$ soit non nul assure que le $\bar{\phi}$ ainsi défini est bien bijectif[8]. Et de plus ce $\bar{\phi}$ vérifie la condition

$$\begin{aligned}\pi\big(\bar{\phi}(z_1,z_2)\big)&=\pi\big((az_1,cz_1)+(bz_2,dz_2)\big) &(24.153a)\\ &=[az_1+bz_2,cz_1+dz_2] &(24.153b)\\ &=\big[a\varphi_0(Z)+b,c\varphi_0(Z)+d\big] &(24.153c)\\ &=\phi\big(\pi(z_1,z_2)\big). &(24.153d)\end{aligned}$$

Nous avons utilisé la notion de classe pour diviser par z_2 et faire apparaitre $\varphi_0(Z)$.

□

Remarque 24.70.
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

$$\bar{\phi}(z)=\frac{az+b}{cz+d} \tag{24.154}$$

avec $ad-cb\neq 0$. Il n'y a pas besoin de traiter séparément le cas $z=\infty$ ou $z=-d/c$.

Définition 24.71.
*Nous aimons tellement l'identification $\varphi_0\colon P(\mathbb{C}^2)\to\hat{\mathbb{C}}$ que nous allons parler d'**homographie** sur $\hat{\mathbb{C}}$ pour les applications $\tilde{\phi}$ de la forme*

$$\tilde{\phi}(z)=\frac{az+b}{cz+d} \tag{24.155}$$

avec $ad-cb\neq 0$.

24.72.
La proposition 24.69 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 $\phi\colon P(\mathbb{C}^2)\to P(\mathbb{C}^2)$.

Proposition 24.73 ([378]).
L'application $h\colon\hat{\mathbb{C}}\to\hat{\mathbb{C}}$ 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 ι 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 24.6.1.5 qui explique qu'une application linéaire sur $\mathbb{C}$ est automatiquement prolongée à $\hat{\mathbb{C}}$ par $f(\infty)=\infty$.

Soit donc l'application

$$h(z)=\frac{az+b}{cz+d}. \tag{24.156}$$

Si $c=0$, alors c'est une application linéaire et la preuve est terminée. Nous supposons que $c\neq 0$. Nous posons

$$l_2(z)=cz+d, \tag{24.157}$$

et ensuite $l_1(z)=\alpha z+\beta$ avec α et β à déterminer. Un peu de calcul :

$$(l_1\circ\iota\circ l_2)(z)=l_1\left(\frac{1}{cz+d}\right)=\frac{\alpha+\beta cz+\beta d}{cz+d}, \tag{24.158}$$

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 vérifier que ces choix donnent le bon résultat. □

8. 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.

24.74.
Vu que les applications linéaires sont des composées d'une translation et d'une rotation-homothétie, et que l'application i 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[9] $z \mapsto 1/\bar{z}$, 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.

Lemme 24.75.
La conjugaison complexe n'est pas une homographie.

Démonstration. Si elle l'était nous aurions des nombres $a, b, c, d \in \mathbb{C}$ tels que $ad - bc \neq 0$ et

$$\frac{az + b}{cz + d} = \bar{z} \tag{24.159}$$

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

$$a = c + d. \tag{24.160}$$

Si $ci + d \neq 0$ Dans ce cas nous pouvons évaluer (24.159) en $z = i$ et avoir $a = -ci + d$. Mais comme nous avions déjà $a = c + d$ nous déduisons $c = 0$. Nous restons donc avec

$$\frac{a}{d} z = \bar{z} \tag{24.161}$$

pour tout z. En prenant $z = 1$ puis $z = i$, il est vite remarqué que cela n'est pas possible.

Si $ci + d = 0$ Nous rappelons que $ad \neq 0$. Nous écrivons l'équation avec $z = -i$ pour trouver

$$\frac{-ai}{-ci + d} = i, \tag{24.162}$$

qui donne immédiatement $a = ci - d$. Nous avons donc les trois équations

$$\begin{cases} c = id & \text{(24.163a)} \\ a = ci - d & \text{(24.163b)} \\ a = c + d. & \text{(24.163c)} \end{cases}$$

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 23.11, 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 24.76 (Inversion de cercles et de droites)**.**
L'inversion dans $\hat{\mathbb{C}}$ envoie

9. Oui, c'est l'inversion de la géométrie hyperbolique, voir 24.6.1.6.

(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.

Droite passant par 0 La façon la plus simple de traiter la droite passant par 0 est de l'écrire sous forme paramétrique :

$$z(t) = te^{i\theta} \tag{24.164}$$

pour θ fixé et $t \in \mathbb{R} \cup \{\infty\}$. En appliquant l'inversion :

$$\iota\big(z(t)\big) = \frac{1}{(te^{i\theta})^*} = \frac{1}{t}e^{i\theta}. \tag{24.165}$$

Notons que les cas particuliers fonctionnent : pour $t = 0$ nous avons le point ∞ et pour $t = \infty$ nous avons 0.

Droite ne passant pas par 0 Une droite ne passant par par $z = 0$ est un ensemble de la forme

$$d(\omega, k) = \{z \in \mathbb{C}^* \text{ tel que } \bar{\omega}z + \omega\bar{z} = k\} \cup \{\infty\} \tag{24.166}$$

avec $k \neq 0$. Étant donné que ι est une bijection et même une involution nous avons $z \in \iota\big(d(\omega, k)\big)$ si et seulement si $\tau(z) \in d(\omega, k)$. L'équation est donc, pour $z \neq 0$:

$$\frac{\bar{\omega}}{z} + \frac{\omega}{\bar{z}} = k. \tag{24.167}$$

Et comme $z \neq 0$ nous pouvons multiplier par $z\bar{z}$ pour trouver $\bar{z}\bar{\omega} + z\omega = kz\bar{z}$. Donc

$$\iota\big(d(\omega, k)\big) = \{z \in \mathbb{C}^* \text{ tel que } \bar{z}\bar{\omega} + z\omega = kz\bar{z}\} \cup \{0\}. \tag{24.168}$$

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 :

$$\iota\big(d(\omega, k)\big) = \{z \in \mathbb{C} \text{ tel que } \bar{\omega}z + \omega\bar{z} = kz\bar{z}\}. \tag{24.169}$$

Cela est l'équation d'un cercle passant par l'origine (définition 24.60).

Cercle ne passant pas par 0 Nous considérons le cercle $C(\omega, r)$ avec $r^2 \neq |\omega|^2$. Il ne contient ni ∞ ni 0 et nous avons alors

$$\iota\big(C(\omega, r)\big) = \{z \in \mathbb{C}^* \text{ tel que } \frac{1}{z\bar{z}} - \bar{\omega}\frac{1}{\bar{z}} - \omega\frac{1}{z} = r^2 - |\omega|^2\}. \tag{24.170}$$

Vu que z n'est jamais nul nous pouvons multiplier l'équation par $z\bar{z}$:

$$\iota\big(C(\omega, r)\big) = \{z \in \mathbb{C} \text{ tel que } \big(|\omega|^2 - r^2\big)z\bar{z} - \bar{\omega}z - \omega\bar{z} = -1\}. \tag{24.171}$$

Le coefficient $|\omega|^2 - r^2$ est non nul par hypothèse et cet ensemble est un cercle par le lemme 24.62. Il ne passe manifestement pas par $z = 0$.

Cercle passant par 0 Le cercle passe par 0, et donc son image par ∞. Nous écrivons alors

$$\iota\big(C(\omega, |\omega|)\big) = \iota\big(C(\omega, |\omega|)\backslash\{0\}\big) \cup \{\infty\}. \tag{24.172}$$

Nous avons

$$C(\omega, |\omega|)\backslash\{0\} = \{z \in \mathbb{C}^* \text{ tel que } z\bar{z} - \bar{\omega}z - \omega\bar{z} = 0\}, \tag{24.173}$$

et un calcul usuel donne

$$\iota\big(C(\omega, |\omega|)\backslash\{0\}\big) = \{z \in \mathbb{C}^* \text{ tel que } 1 - \bar{\omega}z - \omega\bar{z} = 0\}, \tag{24.174}$$

et donc

$$\iota\big(C(\omega, |\omega|)\big) = d(\omega, 1), \tag{24.175}$$

en nous souvenant que le point ∞ est contenu dans $d(\omega, 1)$.

□

Proposition 24.77.
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 23.11 et 24.74 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 24.76 qui donne déjà toutes les réponses.

□

24.78.
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 24.79.
Nous disons qu'une application $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ préserve localement les angles non orientés lorsque

$$\cos\big(df_a(u), df_a(v)\big) = \cos(u, v) \tag{24.176}$$

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 19.53.

Théorème 24.80.
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

$$\iota(z) = \frac{1}{z} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}. \tag{24.177}$$

Nous devons donc étudier la fonction

$$\begin{aligned} F\colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ (x, y) &\mapsto \Big(\frac{x}{r^2}, -\frac{y}{r^2}\Big) \end{aligned} \tag{24.178}$$

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

$$\frac{u \cdot v}{\|u\|\|v\|} = \frac{df_a(u) \cdot df_a(v)}{\|df_a(u)\|\|df_a(v)\|}. \tag{24.179}$$

Pour cela, nous pourrions calculer dF_a et passer en coordonnées polaires[375] 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

$$Au \cdot Av = A^tAu \cdot v \tag{24.180}$$

ainsi que $\|u\| = \sqrt{u \cdot u} = \sqrt{A^tAu \cdot u}$, de telle sorte qu'il devienne urgent de calculer A^tA. Voici le calcul :

```
var('y')

# les fonctions coordonnées
F1(x,y)=x/(x**2+y**2)
F2(x,y)=-y/(x**2+y**2)

# La matrice différentielle :
A=matrix( [  [F1.diff(x).simplify_full(),F1.diff(y).simplify_full()],[F2.diff(x).simplify_full(),F2.diff(y).simplify_full()]  ] )

# Quelqu'un peut expliquer pourquoi ceci ne fonctionne pas ?
# A=matrix( [  [F1.diff(x),F1.diff(y)],[F2.diff(x),F2.diff(y)]  ] ).simplify_full()

# A^tA
S=A.transpose()*A
S=S.simplify_full()      # Mais ça, ça marche !!
print(S)
```

tex/sage/sageSnip009.sage

Le résultat est que

$$A^tA = \begin{pmatrix} \frac{1}{r^4} & 0 \\ 0 & \frac{1}{r^4} \end{pmatrix} = \frac{1}{r^4}\mathrm{Id}. \tag{24.181}$$

La vérification de (24.179) est alors immédiate. □

24.6.3 Birapport

Nous introduisons maintenant quelque chose qui s'appelle le « birapport » et qui n'est a priori pas du tout lié au birapport définit en 24.40.

Définition 24.81 (Birapport dans $\hat{\mathbb{C}}$[313])**.**
*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 $a, b, c \neq \infty$:*

$$[a, b, c, x] = \frac{(a-c)(b-x)}{(b-c)(a-x)}, \tag{24.182}$$

et

$$[\infty, b, c, x] = \frac{b-x}{a-x} \tag{24.183a}$$

$$[a, \infty, c, x] = \frac{a-c}{a-x} \tag{24.183b}$$

$$[a, b, \infty, x] = \frac{b-x}{a-x} \tag{24.183c}$$

24.82.
Notons la « logique » des cas particuliers. Pour le premier, si $a \to \infty$ 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

$$[a, b, \infty, \infty] = 1 \tag{24.184}$$

comme il se doit si nous avons l'intention de ressembler au lemme 24.45.

L'objet « birapport » introduit ici est évidemment lié au birapport sur $P(\mathbb{C}^2)$ définit plus haut. Le lien est la proposition suivante.

Proposition 24.83.
Soit l'application $\varphi_0\colon P(\mathbb{C}^2) \to \hat{\mathbb{C}}$ définie en 24.36. Si $A, B, C, X \in P(\mathbb{C}^2)$ alors

$$[A, B, C, X]_{\varphi_0} = [\varphi_0(A), \varphi_0(B), \varphi_0(C), \varphi_0(X)]. \tag{24.185}$$

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 $a_1, a_2, b_2, b_2, c_1, c_2 \in \mathbb{C}$. Par définition $[A, B, C, X]_{\varphi_0} = (\varphi_0 \circ \phi)(X)$ où $\phi\colon P(\mathbb{C}^2) \to P(\mathbb{C}^2)$ est l'unique homographie telle que

$$\phi[a_1, a_2] = [1, 0] \tag{24.186a}$$
$$\phi[b_1, b_2] = [0, 1] \tag{24.186b}$$
$$\phi[c_1, c_2] = [1, 1] \tag{24.186c}$$

Une des difficultés de cette preuve va être de calculer ce ϕ. D'abord nous pouvons introduire $\tilde{\phi} = \varphi_0 \circ \phi \circ \varphi^{-1}$ qui est obligatoirement (proposition 24.69) de la forme

$$\tilde{\phi}(z) = \frac{\alpha z + \beta}{\gamma a + \delta}. \tag{24.187}$$

Nous allons imposer les relations (24.186) pour déterminer les coefficients α, β, γ et δ.

D'abord

$$(\varphi_0^{-1} \circ \tilde{\phi} \circ \varphi_0)\big([a_1, a_2]\big) = (\varphi_0^{-1} \circ \tilde{\phi})(a_1/a_2) \tag{24.188a}$$
$$= \varphi_0^{-1}\left(\frac{\alpha\frac{a_1}{a_2} + \beta}{\gamma\frac{a_1}{a_2} + \delta}\right) \tag{24.188b}$$
$$= \big[\alpha\frac{a_2}{a_2} + \beta, \gamma\frac{a_1}{a_2} + \delta\big]. \tag{24.188c}$$

Égaler cela à $[1, 0]$ donne

$$\begin{cases} \alpha\dfrac{a_1}{a_2} + \beta \neq 0 & \text{(24.189a)} \\ \gamma\dfrac{a_1}{a_2} + \delta = 0. & \text{(24.189b)} \end{cases}$$

Donc nous avons déjà

$$\tilde{\phi}(z) = \frac{\alpha z + \beta}{\gamma(z - \frac{a_1}{a_2})} = \frac{\alpha z + \beta}{\gamma\big(z - \varphi_0(A)\big)}. \tag{24.190}$$

En y imposant la contrainte $(\varphi^{-1} \circ \tilde{\phi} \circ \varphi_0)([b_1, b_2]) = [0, 1]$ nous trouvons les contraintes

$$\begin{cases} \gamma\big(\varphi_0(B) - \varphi_0(A)\big) \neq 0 & \text{(24.191a)} \\ \beta = -\alpha\varphi_0(B). & \text{(24.191b)} \end{cases}$$

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 (24.185). En substituant :

$$\tilde{\phi}(z) = \frac{\alpha\big(z - \varphi_0(B)\big)}{\gamma\big(z - \varphi_0(A)\big)}. \tag{24.192}$$

La condition pour $[c_1, c_2]$ donne

$$\big[\alpha\big(\varphi_0(C) - \varphi_0(B)\big), \gamma\big(\varphi_0(C) - \varphi_0(A)\big)\big] = [1, 1], \tag{24.193}$$

ce qui donne

$$\alpha\big(\varphi_0(C)-\varphi_0(B)\big)=\gamma\big(\varphi_0(C)-\varphi_0(A)\big). \tag{24.194}$$

Nous avons alors

$$\gamma=\alpha\frac{\varphi_0(C)-\varphi_0(B)}{\varphi_0(C)-\varphi_0(A)}, \tag{24.195}$$

et les α se simplifient dans la formule pour $\tilde{\phi}$:

$$\tilde{\phi}(z)=\frac{\big(z-\varphi_0(B)\big)\big(\varphi_0(C)-\varphi_0(A)\big)}{\big(z-\varphi_0(A)\big)\big(\varphi_0(C)-\varphi_0(B)\big)}. \tag{24.196}$$

Par la proposition 24.69, 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 :

$$\begin{aligned}
[A,B,C,X]_{\varphi_0}&=(\varphi_0\circ\phi)(X) && (24.197a)\\
&=(\varphi_0\circ\varphi_0^{-1}\circ\tilde{\phi}\circ\varphi_0)(X) && (24.197b)\\
&=\tilde{\phi}\big(\varphi_0(X)\big) && (24.197c)\\
&=\frac{\big(\varphi_0(X)-\varphi_0(B)\big)\big(\varphi_0(C)-\varphi_0(A)\big)}{\big(\varphi_0(X)-\varphi_0(A)\big)\big(\varphi_0(C)-\varphi_0(B)\big)} && (24.197d)\\
&=[\varphi_0(A),\varphi_0(B),\varphi_0(C),\varphi_X(A)]. && (24.197e)
\end{aligned}$$

□

Proposition 24.84.
Les homographies de $\hat{\mathbb{C}}$ conservent le birapport.

Démonstration. Ici le mot « homographie » réfère à la définition 24.71 et le birapport à 24.81. Soient $a,b,c,x\in\hat{\mathbb{C}}$ et une homographie $\tilde{\phi}\colon\hat{\mathbb{C}}\to\hat{\mathbb{C}}$. Il existe une homographie $\phi\colon P(\mathbb{C}^2)\to P(\mathbb{C}^2)$ telle que $\tilde{\phi}=\varphi_0\circ\phi\varphi_0^{-1}$. Alors

$$\begin{aligned}
\big[\tilde{\phi}(a),\tilde{\phi}(b),\tilde{\phi}(c),\tilde{\phi}(x)\big]&=\big[(\phi\circ\varphi_0^{-1})(a),(\phi\circ\varphi_0^{-1})(b),(\phi\circ\varphi_0^{-1})(c),(\phi\circ\varphi_0^{-1})(x),\big]_{\varphi_0} && (24.198a)\\
&=\big[\varphi_0^{-1}(a),\varphi_0^{-1}(b),\varphi_0^{-1}(c),\varphi_0^{-1}(x),\big]_{\varphi_0} && (24.198b)\\
&=[a,b,c,x]. && (24.198c)
\end{aligned}$$

Justifications :

— Identification des birapports sur $P(\mathbb{C}^2)$ et sur $\hat{\mathbb{C}}$, proposition 24.83.
— Invariance du birapport sour les homographies (dans $P(\mathbb{C}^2)$), proposition 24.47.

□

Proposition 24.85.
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 $[a,b,c,x]\in\hat{\mathbb{R}}$.

Démonstration. Nous allons faire plusieurs cas. Mais dans tous les cas vous pouvez relire la définition des angles orientés 19.146 et la partie sur les angles dans les nombres complexes 19.10.11.

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 (24.182) est

$$[\overrightarrow{ca},\overrightarrow{cb}]+[\overrightarrow{xb},\overrightarrow{xa}]. \tag{24.199}$$

Le birapport est réel si et seulement si

$$[\overrightarrow{ca},\overrightarrow{cb}]-[\overrightarrow{xa},\overrightarrow{xb}]\in[0]_\pi. \tag{24.200}$$

À gauche nous avons une classe modulo 2π et à droite une classe modulo π. 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 :

$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] = [0]_\pi, \tag{24.201a}$$

$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] \in [0]_\pi, \tag{24.201b}$$

$$[\overrightarrow{ca}, \overrightarrow{cb}] - [\overrightarrow{xa}, \overrightarrow{xb}] \subset [0]_\pi. \tag{24.201c}$$

Nous avons donc que le birapport est réelle si et seulement si la condition (24.201c) est vérifiée. D'après le théorème 19.154, cette dernière condition est équivalente à dire que les points a, b, c et x sont alignés.

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[10]. 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 ∞. Si l'un de a, b ou c est ∞ et si x l'est aussi, alors, comme ∞ 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 ∞.

Si $a = \infty$ Le birapport est alors

$$[\infty, b, c, x] = \frac{b - x}{a - x}, \tag{24.202}$$

qui est un nombre a 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

$$\arg\left(\frac{b - x}{a - x}\right) \in [\overrightarrow{xa}, \overrightarrow{xb}]. \tag{24.203}$$

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. □

24.6.4 Division harmonique

Définition 24.86.
*Nous disons que les éléments a, b, c et x de $\hat{\mathbb{C}}$ sont en **division harmonique** lorsque $[a, b, c, x] = -1$.*

24.87.
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 une 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.

10. 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.

Lemme 24.88 ([379, 380]).
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

$$x = (mc) \cap \mathcal{C} \tag{24.204}$$

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.

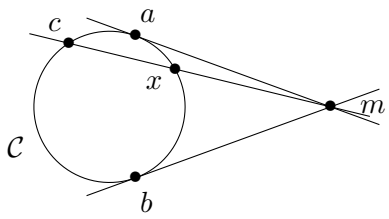
Démonstration. Nous séparons les cas suivant que $m = \infty$ ou non.

$\underline{m = \infty}$ 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 24.84), et les rotations, dilatations et translations sont des homographies (voir 24.74).

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 :

$$[a, b, c, x] = \frac{(i - e^{i\theta})(-i - e^{i(\pi-\theta)})}{(-i - e^{i\theta})(i - e^{i(\pi-\theta)})} = -1. \tag{24.205}$$

$\underline{m \neq \infty}$ 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[11].

Nous nommons $A = (am)$, $B = (bm)$, $C = (cm)$ et $\mathcal{C}$, le cercle. Nous allons maintenant faire intensément usage de la proposition 24.76. 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 ∞ (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 $\tilde{\phi}(\mathcal{C})$.

Par tangence, la droite A et le cercle $\mathcal{C}$ se coupent en un seul point (a). Donc $\tilde{\phi}(A)$ coupe $\tilde{\phi}(\mathcal{C})$ 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 $\tilde{\phi}(\mathcal{C})$ et se coupent en exactement deux points distincts (qui sont donc du même côté de la droite).

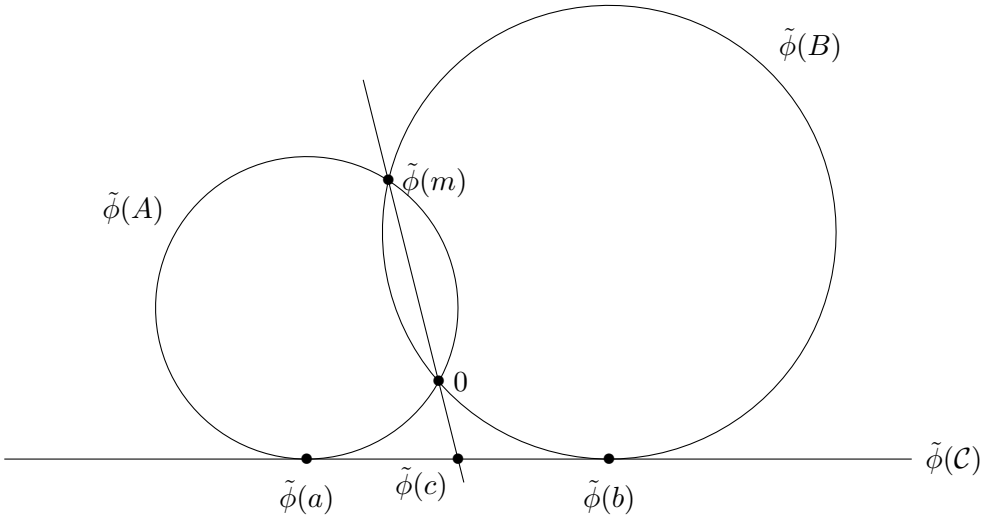
Nous nous intéressons à la droite $\tilde{\phi}(C)$. C'est une droite parce que c'est l'image d'une droite passant par 0. Elle passe par 0, par $\tilde{\phi}(c)$ et $\tilde{\phi}(m)$. Le fait qu'elle passe par 0 et $\tilde{\phi}(m)$ fait que c'est la droite passant par les deux intersections des cercles. Vu que $c \in \mathcal{C} \cap C$, le point d'intersection $\tilde{\phi}(C) \cap \tilde{\phi}(\mathcal{C})$ est $\tilde{\phi}(c)$.

11. 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 θ et σ. Ensuite le calcul du birapport est explicite.

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 :

$$k = d\big(\tilde{\phi}(c), 0\big) \times d\big(\tilde{\phi}(c), \tilde{\phi}(m)\big). \tag{24.206}$$

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 [12] :



Mais la droite passant par $\tilde{\phi}(a)$ et $\tilde{\phi}(c)$ (qui est tangente au cercle) permet également de calculer cette puissance :

$$k = d\big(\tilde{\phi}(a), \tilde{\phi}(c)\big). \tag{24.207}$$

Idem pour la puissance par rapport à l'autre cercle :

$$k = d\big(\tilde{\phi}(a), \tilde{\phi}(b)\big). \tag{24.208}$$

Nous en déduisons que $\tilde{\phi}(c)$ est le milieu entre $\tilde{\phi}(a)$ et $\tilde{\phi}(b)$.
Du coup

$$\big[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \tilde{\phi}(x)\big] = \big[\tilde{\phi}(a), \tilde{\phi}(b), \tilde{\phi}(c), \infty\big] = -1 \tag{24.209}$$

en vertu de ce que nous avons raconté en 24.87.

□

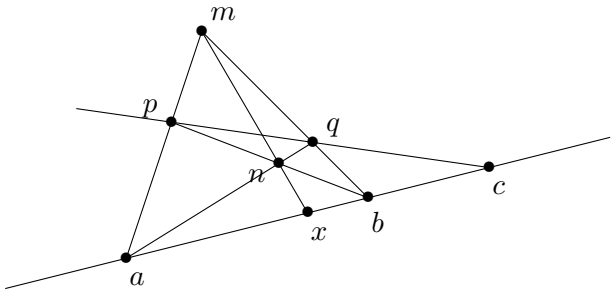
Lemme 24.89 ([381]).
Soit $a, b, c \in \mathbb{C}$ colinéaires. Soit m un point hors de cette droite. Nous considérons une droite issue de c coupant $[ma]$ en p et $[mb]$ en q.

Nous construisons $n = (aq) \cap (pb)$ et finalement $x = (mn) \cap (ab)$.

À la fin nous avons

$$[a, b, c, x] = -1. \tag{24.210}$$

Démonstration. Commençons par un dessin de la situation :



12. 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 $\big(\tilde{\phi}(m)\tilde{\phi}(c)\big)$ n'arrive pas perpendiculairement à $\tilde{\phi}(\mathcal{C})$?

Les points a, b et m ne sont pas alignés, et nous pouvons les utiliser comme repère barycentrique (voir 10.47 pour savoir en deux mots ce que c'est). Nous nommons (α, β, γ) les coordonnées de n dans ce système, c'est-à-dire que

$$\alpha\overrightarrow{na} + \beta\overrightarrow{nb} + \gamma\overrightarrow{nm} = 0. \tag{24.211}$$

Dans notre contexte, nous pouvons voir le vecteur $\overrightarrow{st}$ comme une façon d'écrire le nombre $t - s$. Par la proposition 10.52 nous savons les coordonnées barycentriques de p, x et q en regardant le triangle acb. Voici les coordonnées et les relations qu'elles signifient :

$$n = (\alpha, \beta, \gamma), \qquad \alpha\overrightarrow{na} + \beta\overrightarrow{nb} + \gamma\overrightarrow{nm} = 0 \tag{24.212a}$$
$$p = (\alpha, 0, \gamma), \qquad \alpha\overrightarrow{pa} + \gamma\overrightarrow{pm} = 0 \tag{24.212b}$$
$$q = (0, \beta, \gamma), \qquad \beta\overrightarrow{qb} + \gamma\overrightarrow{qm} = 0 \tag{24.212c}$$
$$x = (\alpha, \beta, 0). \qquad \alpha\overrightarrow{xa} + \beta\overrightarrow{xb} = 0. \tag{24.212d}$$

Nous voudrions maintenant voir les coordonnées de c. Nous posons $c = (\lambda, \mu, \sigma)$:

$$\lambda\overrightarrow{ca} + \mu\overrightarrow{cb} + \sigma\overrightarrow{cm} = 0. \tag{24.213}$$

Mais a, b et c sont alignés, donc $\overrightarrow{ca}$ et $\overrightarrow{cb}$ sont colinéaires, alors que $\overrightarrow{cm}$ n'est pas aligné avec les deux autres. L'annulation (24.213) demande donc l'annulation séparément

$$\begin{cases} \sigma\overrightarrow{cm} = 0 & \text{(24.214a)} \\ \lambda\overrightarrow{ca} + \mu\overrightarrow{cb} = 0. & \text{(24.214b)} \end{cases}$$

Nous en déduisons que $\sigma = 0$ et aussi que λ et μ 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 apparaitre les vecteurs $\overrightarrow{pq}$, $\overrightarrow{pc}$, $\overrightarrow{pc}$. Vu le dessin et les relations disponibles (24.212) le mieux est d'utiliser les relations de Chasles (proposition 10.3) pour faire $\overrightarrow{ca} = \overrightarrow{cp} + \overrightarrow{pq}$ et $\overrightarrow{cb} = \overrightarrow{cq} + \overrightarrow{qb}$. La relation (24.214b) devient :

$$\overrightarrow{cp} + \overrightarrow{pa} + \mu(\overrightarrow{cq} + \overrightarrow{qb}) = 0. \tag{24.215}$$

Les vecteurs $\overrightarrow{cp}$ et $\overrightarrow{cq}$ sont alignés, donc nous les écrivons ensemble. Les vecteurs $\overrightarrow{pa}$ et $\overrightarrow{qb}$ se transforment en utilisant les relations (24.212) :

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}\overrightarrow{pm} - \mu\frac{\gamma}{\beta}\overrightarrow{qm}. \tag{24.216}$$

Enfin nous voulons faire la somme du terme $\overrightarrow{pm}$ avec le terme $\overrightarrow{qm}$. D'abord on change le signe :

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}\overrightarrow{pm} + \frac{\mu\gamma}{\beta}\overrightarrow{mq} \tag{24.217}$$

ensuite nous écrivons

$$\frac{\mu\gamma}{\beta} = \frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} - \frac{\gamma}{\alpha}, \tag{24.218}$$

et

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}(\overrightarrow{pm} + \overrightarrow{mq}) + \left(\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha}\right)\overrightarrow{mq} = 0. \tag{24.219}$$

Tout cela pour

$$\overrightarrow{cp} + \mu\overrightarrow{cq} - \frac{\gamma}{\alpha}\overrightarrow{pq} + \left(\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha}\right)\overrightarrow{mq} = 0. \tag{24.220}$$

Le dernier terme n'est pas colinéaire aux deux premiers et s'annule donc séparément :

$$\frac{\mu\gamma}{\beta} + \frac{\gamma}{\alpha} = 0. \tag{24.221}$$

Cela donne $\mu = -\beta/\alpha$.

Au finale nous avons

$$\overrightarrow{ca} - \frac{\beta}{\alpha}\overrightarrow{cb} = 0 \tag{24.222}$$

et donc

$$\alpha\overrightarrow{ca} - \beta\overrightarrow{cb} = 0, \tag{24.223}$$

ce qui donne les coordonnées $(\alpha, -\beta, 0)$ pour le point c.

Vu que nous somme dans l'espace vectoriel $\mathbb{C}$, ce que nous notons $\overrightarrow{AB}$ n'est rien d'autre que la différence $B - A$ dans $\mathbb{C}$[13]. La relation (24.223) signifie donc

$$\alpha(a - c) = \beta(b - c). \tag{24.224}$$

Nous avons alors :

$$\frac{a - c}{b - c} = \frac{\beta}{\alpha}. \tag{24.225}$$

Par ailleurs, la relation (24.212d) à propos des coordonnées de x donne

$$\frac{a - x}{b - x} = -\frac{\beta}{\alpha}. \tag{24.226}$$

En égalisant les deux valeurs de β/α nous trouvons :

$$\frac{a - c}{b - c} = \frac{x - a}{b - x}, \tag{24.227}$$

ce qui donne (via un petit jeu de signes)

$$\frac{(c - a)(x - b)}{(c - b)(x - a)} = -1. \tag{24.228}$$

C'est cela que nous voulions. □

24.6.5 Groupe circulaire

Nous avons vu que les homographies présent l'ensemble des cercles et droites. Nous pouvons nous demander quel est le groupe maximum préservant l'ensemble des cercles et droites.

Définition 24.90.
*Le **groupe circulaire** de $\mathbb{C}$ est le groupe de transformations de $\hat{\mathbb{C}}$ engendré par les homographies*[14] *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 eC.*

Plusieurs remarques à propos de cette définition.

(1) Vu le lemme 24.75, 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.

(2) Vous vous souvenez de la définition d'un sous-groupe engendré ? C'est la définition 2.6.

Lemme 24.91.
Soit une application $\alpha\colon \mathbb{C} \to \mathbb{C}$ 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 α est un automorphisme de corps[15].

13. Les mauvaise 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.
14. Homographie de $\hat{\mathbb{C}}$: définition 24.71.
15. Définition 1.42.

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 α préserve les divisions harmoniques nous avons équivalence entre les affirmations suivantes :

$$c = \frac{a+b}{2} \tag{24.229a}$$
$$[a, b, c, \omega] = -1 \tag{24.229b}$$
$$[\alpha(a), \alpha(b), \alpha(c), \alpha(\infty)] = -1 \tag{24.229c}$$
$$\frac{\alpha(a) + \alpha(b)}{2} = \alpha(c). \tag{24.229d}$$

Donc α préserve les milieux : pour tout $a, b \in \mathbb{C}$ nous avons

$$\alpha\left(\frac{a+b}{2}\right) = \frac{\alpha(a) + \alpha(b)}{2}. \tag{24.230}$$

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 (24.230) :

$$\frac{\alpha(a) + \alpha(b)}{2} = \alpha\left(\frac{a+b}{2}\right) = \frac{\alpha(a+b)}{2}. \tag{24.231}$$

Cela démontre déjà que

$$\alpha(a + b) = \alpha(a) + \alpha(b). \tag{24.232}$$

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 ± 1 nous avons

$$[a, -a, a^2, 1] = \frac{(a - a^2)(-a - 1)}{(-a - a^2)(a - 1)} = -1. \tag{24.233}$$

Et en prenant $\alpha(a)$ en guise de a nous avons aussi

$$[\alpha(a), -\alpha(a), \alpha(a)^2, \alpha(1)] = -1. \tag{24.234}$$

Vu que α préserve les divisions harmoniques, l'équation (24.233) donne aussi

$$\big[\alpha(a), \alpha(-a), \alpha(a^2), \alpha(1)\big] = -1, \tag{24.235}$$

c'est-à-dire

$$\big[\alpha(a), -\alpha(a), \alpha(a^2), 1\big] = -1. \tag{24.236}$$

Comparant (24.234) avec (24.236) et en tenant compte de l'unicité du birapport [16] nous avons

$$\alpha(a^2) = \alpha(a)^2. \tag{24.237}$$

Avec cela nous pouvons y aller en remarquant que

$$ab = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2. \tag{24.238}$$

Nous appliquons α à cette dernière équations en tenant compte de ce que nous savons déjà

$$\alpha(ab) = \left(\frac{\alpha(a) + \alpha(b)}{2}\right)^2 - \left(\frac{\alpha(a) - \alpha(b)}{2}\right)^2 = \alpha(a)\alpha(b). \tag{24.239}$$

□

16. 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 24.46(2).

Théorème 24.92 ([381, 382]).
Le groupe circulaire de $\mathbb{C}$ est le groupe des bijections $\hat{\mathbb{C}} \to \hat{\mathbb{C}}$ 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 24.77. Et la conjugaison complexe aussi.

Soit une bijection $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ préservant les cercles-droites. Nous supposons dans un premier temps que $f(0) = 0$, $f(1) = 1$ et $f(\infty) = \infty$.

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 ∞, 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.

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 24.89. 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 $x = (mn) \cap (ac)$.

Vues les propriétés de f (en particulier c'est une bijection qui respecte les intersectons, tangences, cercles et droites). Le point $f(m)$ est hors de la droite $\big(f(a)f(b)\big)$. 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

$$f(n) = \big(p(p)f(b)\big) \cap \big(f(q)f(a)\big) \tag{24.240}$$

et aussi

$$f(x) = \big(p(m)f(n)\big) \cap \big(f(a)f(c)\big) \tag{24.241}$$

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 24.88 nous donne le même résultat.

f est un automorphisme du corps $\mathbb{C}$ C'est le lemme 24.91.

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.

Pour f plus générale Nous ne supposons plus que f fixe 0, 1 et ∞. 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 [17] $\phi\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ telle que $\phi(1) = f^{-1}(1)$, $\phi(0) = f^{-1}(0)$ et $\phi(\infty) = f^{-1}(\infty)$. Dans ce cas l'application

$$g = f \circ \phi \tag{24.242}$$

vérifie $g(1) = 1$, $g(0) = 1$ 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

$$f = g \circ \phi^{-1} \tag{24.243}$$

est la composée d'une homographie avec soit l'identité soit la conjugaison complexe. Elle est donc dans le groupe circulaire.

□

17. 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 corollaire 24.35 avec l'application φ_0 et la définition 24.71 et 24.72.

24.6.6 Action du groupe modulaire

Le **demi-plan de Poincaré** est l'ensemble

$$P = \{z \in \mathbb{C} \text{ tel que } \Im(z) > 0\}. \tag{24.244}$$

Le **groupe modulaire** est le quotient de groupes

$$\mathrm{PSL}(2,\mathbb{Z}) = \frac{\mathrm{SL}(2,\mathbb{Z})}{\mathbb{Z}_2}. \tag{24.245}$$

Ce sont donc les matrices au signe près de la forme

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{24.246}$$

où a, b, c et d sont entiers tels que $ad - cb = 1$.

Théorème 24.93 ([383]).
Le groupe modulaire agit fidèlement (définition 2.46) sur le demi-plan de Poincaré par

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} * z = \frac{az + b}{cz + d}. \tag{24.247}$$

L'ensemble $D = D_1 \cup D_2$ avec

$$D_1 = \{z \in P \text{ tel que } |z| > 1, -\frac{1}{2} \leqslant \Re(z) < \frac{1}{2}\} \tag{24.248a}$$

$$D_2 = \{z \in P \text{ tel que } |z| = 1, -\frac{1}{2} \leqslant \Re(z) \leqslant 0\} \tag{24.248b}$$

est un domaine fondamental (définition 2.53) de cette action.

De plus si nous notons

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{24.249}$$

alors pour tout $z \in P$, il existe $A \in \mathrm{gr}(S,T)$ telle que $A * z \in D$.

Démonstration. Nous divisions la preuve en plusieurs étapes.

Bien définie D'abord il faut remarquer que l'action (24.247) est bien définie par rapport au quotient : $A * z = (-A) * z$. La vérification est immédiate.

Interne Montrons que si $A \in \mathrm{PSL}(2,\mathbb{Z})$ et $z \in P$ alors $A * z \in P$. Nous avons

$$A * z = \frac{az + b}{cz + d} = \frac{(az + b)(c\bar{z} + d)}{|cz + d|^2} = \frac{a|z|c + azd + bc\bar{z} + bd}{|cz + d|^2}, \tag{24.250}$$

et donc en décomposant $z = \Re(z) + i\Im(z)$,

$$\Im(A * z) = \Im\left(\frac{azd + bc\bar{z}}{|cz + d|^2}\right) = \frac{ad - bc}{|cz + d|^2}\Im(z) = \frac{\Im(z)}{|cz + d|^2} \tag{24.251}$$

où nous avons tenu compte de $ad - bc = 1$. Donc l'action respecte la (stricte) positivité de la partie imaginaire.

Action Nous vérifions maintenant que la formule donne bien une action : $A*(B*z)=(AB)*z$. Cela est un bon calcul :

$$A*(B*z)=A*\left(\frac{a'z+b}{c'z+d'}\right) \tag{24.252a}$$
$$=\frac{a\left(\frac{a'z+b}{c'z+d'}\right)+b}{c\left(\frac{a'z+b}{c'z+d'}\right)+d} \tag{24.252b}$$
$$=\frac{a(a'z+b')+b(c'z+d')}{c(a'z+b')+d(c'z+d')} \tag{24.252c}$$
$$=\frac{(aa'+bc')z+(ab'+bd')}{(ca'+dc')z+(cb'+dd')} \tag{24.252d}$$
$$=\begin{pmatrix}aa'+bc' & ab'+bd'\\ a'c+dc' & cb'+dd'\end{pmatrix}*z \tag{24.252e}$$
$$=(AB)*z. \tag{24.252f}$$

Fidèle Soit $A\in \mathrm{PSL}(2,\mathbb{Z})$ tel que pour tout $z\in P$ nous ayons

$$\frac{az+b}{cz+d}=z. \tag{24.253}$$

Alors nous avons

$$cz^2+(d-a)z+b=0. \tag{24.254}$$

Cela est donc un polynôme en z qui s'annule sur un ouvert [18] (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 $\epsilon>0$) :

$$-c\epsilon^2+\epsilon(d-a)i+b=0 \tag{24.255}$$

pour tout ϵ. Le fait d'avoir $c\epsilon^2=b$ pour tout ϵ implique que $c=b=0$. Donc A est de la forme

$$A=\begin{pmatrix}a & 0\\ 0 & d\end{pmatrix}, \tag{24.256}$$

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 $\mathrm{PSL}(2,\mathbb{Z})$.

Les orbites intersectent D Soit $z\in P$. Nous devons trouver $A\in\mathrm{PSL}(2,\mathbb{Z})$ tel que $A*z\in D$. Nous savons déjà que

$$\Im(A*z)=\frac{\Im(z)}{|cz+d|^2}. \tag{24.257}$$

Nous notons $\mathcal{O}_z$ l'orbite de z sous le groupe modulaire et nous posons

$$I_z=\{\Im(u)\text{ tel que }u\in\mathcal{O}_z\}=\{\Im(A*z)\text{ tel que }A\in\mathrm{PSL}(2,\mathbb{Z})\}, \tag{24.258}$$

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,) z donné, prendre qu'un nombre fini de valeurs plus grandes que $\Im(z)$ [19]. Nous cherchons donc les couples $(c,d)\in\mathbb{Z}^2$ tels que $|cz+d|<1$.

18. 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.

19. 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 $|cz+d|\geqslant|c\Im(z)|$, 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

$$\Re(cz+d)=c\Re(z)+d, \tag{24.259}$$

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 $A_1\in\mathrm{PSL}(2,\mathbb{Z})$ tel que $\Im(A_1*z)=\max I_z$. Nous notons $z_1=A_1*z$, et que nous n'avons a priori pas l'unicité. Nous allons maintenant agir sur z_1 avec l'élément

$$T=\begin{pmatrix}1&1\\0&1\end{pmatrix} \tag{24.260}$$

pour ramener z_1 dans le domaine D. Si $u\in P$ nous avons $T*u=u+1$ et donc

$$T^n*u=u+n. \tag{24.261}$$

Vu que D est de largeur 1, il existe un n (éventuellement négatif) tel que

$$\Re(T^n*z_1)\in[-\frac{1}{2},\frac{1}{2}[. \tag{24.262}$$

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

$$S=\begin{pmatrix}0&-1\\1&0\end{pmatrix} \tag{24.263}$$

qui fait

$$\Im(S*z)=\frac{\Im z}{|z|^2}. \tag{24.264}$$

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 $|z_2|\geqslant 1$. Nous en déduisons que $|z_2|\geqslant 1$.

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 $2\cos(\theta)\leqslant 1$ ou encore $|\Re(z_2)|\leqslant\frac{1}{2}$. Donc si $\Re(z_2)\leqslant 0$ 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 $S*(\rho e^{-i\theta})=-\frac{1}{\rho}e^{-i\theta}$.

Unicité Nous voulons à présent montrer que si $z\in D$, alors $A*z$ n'est plus dans D (sauf si $A=\pm\mathbb{1}$). Nous supposons que $z\in D$ et $A\in\mathrm{PSL}(2,\mathbb{Z})$ soient tels que $A*z\in D$, et nous prouvons qu'alors soit nous arrivons à une contradiction soit nous arrivons à $A=\mathbb{1}$. Pour cela nous allons décomposer en de nombreux cas.

(1) Nous commençons par $\Im(A*z)\geqslant\Im(z)$. Dans ce cas nous avons $|cz+d|\leqslant 1$ et en particulier $|c||\Im(z)|\leqslant 1$. 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 $|c|\leqslant 2/\sqrt{3}$. Vu que c doit être entier, nous avons trois cas : $c=-1,0,1$.

(1a) 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 $\mathrm{PSL}(2,\mathbb{Z})$). La matrice A doit alors être de la forme

$$A=\begin{pmatrix}1&b\\0&1\end{pmatrix} \tag{24.265}$$

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.

(1b) Soit $c = 1$. Alors la condition $|cz + d| \leqslant 1$ nous donne trois possibilités[20] : $d = -1, 0, 1$.

i. Si $d = -1$, alors nous devons avoir $|z-1| \leqslant 1$. 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 $|z - 1| \leqslant 1$.

ii. Si $d = 1$, alors (et c'est maintenant que la dissymétrie de D intervient) nous avons le point

$$z = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \tag{24.266}$$

qui est dans D et qui vérifie $|z + 1| \leqslant 1$. Voyons à quoi ressemble la matrice A dans ce cas. Son déterminant est $a - b = 1$. Nous écrivons donc

$$A = \begin{pmatrix} b+1 & b \\ 1 & 1 \end{pmatrix}, \tag{24.267}$$

et en tenant compte du fait que $z\bar{z} = |z + 1| = 1$, nous calculons

$$A * z = \frac{(b+1)z + b}{z+1} \tag{24.268a}$$

$$= \frac{(bz + z + b)(\bar{z} + 1)}{|z+1|^2} \tag{24.268b}$$

$$= z + b + 1. \tag{24.268c}$$

La seule façon de ne pas quitter D est d'avoir $b = -1$, mais alors nous avons

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \tag{24.269}$$

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

$$z_0 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i. \tag{24.270}$$

C'est un point de qui vérifie $z_0 = A * z_0$ pour un élément non trivial A de $\mathrm{PSL}(2, \mathbb{Z})$. 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.

iii. Le cas $d = 0$ nous fait écrire $1 = \det A = -b$, donc $b = -1$ et

$$A = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}. \tag{24.271}$$

Nous avons alors $A * z = a - \frac{1}{z}$. De plus la condition $|z| \leqslant 1$ revient à $|z = 1|$. Pour les nombres complexes de module 1, l'opération $z \to -1/z$ 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

$$A = \begin{pmatrix} a & b \\ -1 & d \end{pmatrix}, \tag{24.272}$$

et nous revenons au cas $c = 1$ en prenant $-A$ au lieu de A.

20. Je ne rigolais pas quand je disais qu'on allait avoir de nombreux cas.

(2) Nous passons au cas $\Im(A*z)<\Im(z)$. Nous récrivons cette condition avec

$$\Im(A*z)<\Im\big(A^{-1}*(A*z)\big). \tag{24.273}$$

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

$$\Im(z')<\Im(A^{-1}*z'). \tag{24.274}$$

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.

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 $A\in\mathrm{PSL}(2,\mathbb{Z})$ 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}.

□

Corollaire 24.94 ([384])**.**
Les matrices S et T génèrent le groupe modulaire au sens où toute matrice de $\mathrm{PSL}(2,\mathbb{Z})$ *s'écrit comme*

$$T^{m_1}S^{p_1}\dots T^{m_k}S^{p_k} \tag{24.275}$$

pour un certain k et des nombres $m_i,p_i\in\mathbb{Z}$. Autrement dit, $\mathrm{PSL}(2,\mathbb{Z})=\mathrm{gr}(S,T)$.

Démonstration. Soit z, un point de D autre que z_0. Alors si $A\in\mathrm{PSL}(2,\mathbb{Z})$ est non trivial nous avons $A*z$ hors de D. Du coup, comme vu dans la démonstration du théorème 24.93, il existe $B\in\mathrm{gr}(S,T)$ tel que $B*(A*z)\in D$. Vu que D ne contient qu'un seul point de chaque orbite, nous avons

$$B*A*z=z, \tag{24.276}$$

et donc $BA=\pm\mathbb{1}$, ce qui prouve que[21] $A=B^{-1}$, c'est-à-dire que $A\in\mathrm{gr}(S,T)$. □

21. Dans $\mathrm{PSL}(2,\mathbb{Z})$, nous n'avons pas besoin de mettre $\pm$ parce qu'il est compris dans la définition.

Chapitre 25

Analyse vectorielle

25.1 Le théorème de Green

Soit un champ de vecteurs

$$F(x,y,z) = \begin{pmatrix} F_1(x,y,z) \\ F_2(x,y,z) \\ F_2(x,y,z) \end{pmatrix} \tag{25.1}$$

et un chemin $\sigma\colon [a,b] \to \mathbb{R}^3$ donné par

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}. \tag{25.2}$$

Nous avons défini la circulation de F le long de σ par

$$\begin{aligned} \int_\sigma F \cdot d\sigma &= \int_a^b F\big(\sigma(t)\big) \cdot \sigma'(t)dt \\ &= \int_a^b \Big[F_1\big(\sigma(t)\big)x'(t) + F_2\big(\sigma(t)\big)y'(y) + F_3\big(\sigma(t)\big)z'(t)\Big]dt \\ &= \int_\sigma F_1dx + F_2dy + F_3dz. \end{aligned} \tag{25.3}$$

La dernière ligne est juste une notation compacte[1]. 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

$$\int_\sigma f dy + g dz. \tag{25.4}$$

Cela signifie

$$\int_a^b \Big[f\big(\sigma(t)\big)y'(t) + g\big(\sigma(t)\big)z'(t)\Big]dt. \tag{25.5}$$

Soit D une région du plan et σ, son contour que nous prenons, par convention[2], dans l'orientation trigonométrique, comme indiqué sur la figure 25.1. 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 aurai 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...

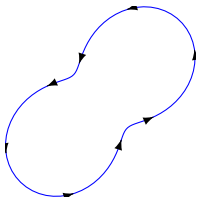


FIGURE 25.1 – Un contour avec son ortientation.

Théorème 25.1 (Théorème de Green)**.**
Soient $P, Q\colon D \to \mathbb{R}$ deux fonctions de classe C^1. Alors

$$\int_{\partial D} P dx + Q dy = \int_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dxdy. \tag{25.6}$$

Pour rappel, l'intégrale du membre de gauche signifie

$$\int_a^b \Big[P\big(\sigma(t)\big)\sigma'_x(t) + Q\big(\sigma(t)\big)\sigma'_y(t) \Big] dt. \tag{25.7}$$

Ce n'est d'ailleurs rien d'autre que l'intégrale du champ de vecteurs $\begin{pmatrix} P \\ Q \end{pmatrix}$.

Corollaire 25.2.
L'aire du domaine D est donnée par

$$A = \frac{1}{2} \int_{\partial D} (x dy - y dx). \tag{25.8}$$

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

$$\begin{aligned} \int_{\partial D} -y dx + x dy &= \int_D \left(\frac{\partial x}{\partial x} - \frac{\partial(-y)}{\partial y} \right) dxdy \\ &= \int_D 2\, dxdy. \end{aligned} \tag{25.9}$$

La dernière ligne est bien le double de la surface. □

Exemple 25.3
Calculons (encore une fois) l'aire du disque de rayon R. Il s'agit de calculer l'intégrale

$$I = \frac{1}{2} \int_\sigma (x dt - y dx) \tag{25.10}$$

où σ est le cercle donné par

$$\sigma(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} R\cos(t) \\ R\sin(t) \end{pmatrix} \tag{25.11}$$

Le calcul est

$$\begin{aligned} I &= \frac{1}{2} \int_0^{2\pi} \underbrace{R\cos(\theta)}_{x} \underbrace{R\cos(\theta)}_{y'} - \underbrace{R\sin(\theta)}_{y} \underbrace{\big(-R\sin(\theta)\big)}_{x'} d\theta \\ &= \frac{R^2}{2} \int_0^\pi d\theta \\ &= \pi R^2. \end{aligned} \tag{25.12}$$

△

Exemple 25.4
Calculons l'aire de l'ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}\leqslant 1 \tag{25.13}$$

dont le bord est donné par

$$\begin{cases} x(t)=a\cos(t) & (25.14a)\\ y(t)=b\sin(t). & (25.14b)\end{cases}$$

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

$$\frac{1}{2}\int_{\partial D}(xdy-ydx)=\frac{1}{2}\int_0^{2\pi}ab=\pi ab. \tag{25.15}$$

△

Le théorème de Green peut être mis sous une autre forme.

Théorème 25.5 (Théorème de Green, forme vectorielle)**.**
Si G est un champ de vecteurs sur D, nous avons

$$\int_{\partial D}G\cdot d\sigma=\int_D(\nabla\times G)\cdot dS \tag{25.16}$$

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

$$\frac{\partial F_2}{\partial x}-\frac{\partial F_1}{\partial y}. \tag{25.17}$$

Cela est bien le membre de droite de l'équation (25.6). Le membre de gauche de cette dernière est bien le membre de gauche de (25.16). □

Exemple 25.6
Soit le champ de vecteurs $F(x,y)=\begin{pmatrix}xy^2\\ y+x\end{pmatrix}$, et soit à calculer

$$\int_D\nabla\times F\cdot dS \tag{25.18}$$

où D est la région comprise entre les courbes $y=x^2$ et $y=x$ pour $x\geqslant 0$ (voir la figure 25.2).

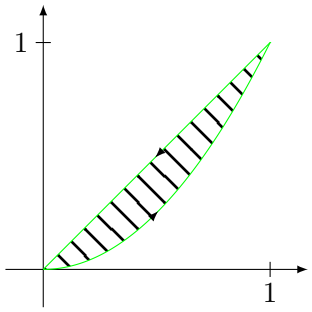


FIGURE 25.2 – Le contour d'intégration pour l'exemple 25.6.

Nous pouvons calculer cette intégrale directement en calculant le rotationnel de F :

$$\nabla \times F = \begin{pmatrix} 0 \\ 0 \\ 1-2xy \end{pmatrix}. \tag{25.19}$$

Par conséquent l'intégrale à effectuer est

$$I = \int_0^1 dx \int_{x^2}^x (1-2xy)dy = \frac{1}{12}. \tag{25.20}$$

```
----------------------------------------------------------------------
| 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 ∂D :

$$I = \int_{\partial D} F \cdot \sigma. \tag{25.21}$$

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

$$\sigma_1(t) = (t, t^2) \tag{25.22}$$

et

$$\sigma_2(t) = (1-t, 1-t). \tag{25.23}$$

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 :

$$\begin{aligned} \sigma_1(t) &= (t, t^2) \\ \sigma_1'(t) &= (1, 2t) \\ F\big(\sigma_1(t)\big) &= \begin{pmatrix} t^5 \\ t+t^2 \end{pmatrix}, \end{aligned} \tag{25.24}$$

et par conséquent

$$F\big(\sigma_1(t)\big) \cdot \sigma_1'(t) = t^5 + 2t^2 + 2t^3, \tag{25.25}$$

et le premier morceau de la circulation vaut

$$\int_{\sigma_1} F \cdot d\sigma_1 = \int_0^1 t^5 + 2t^2 + 2t^3 = \frac{4}{3}. \tag{25.26}$$

Pour le second chemin :

$$\begin{aligned} \sigma_2(t) &= (1-t, 1-t) \\ \sigma_2'(t) &= (-1, -1) \\ F\big(\sigma_2(t)\big) &= \begin{pmatrix} (1-t)^3 \\ 2(1-t) \end{pmatrix}. \end{aligned} \tag{25.27}$$

Par conséquent

$$F\big(\sigma_2(t)\big) \cdot \sigma_2(t) = -(1-t)^2 - 2(1-t). \tag{25.28}$$

Le second morceau de la circulation est par conséquent

$$\int_0^1 -(1-t)^2 - 2(1-t)dt = -\frac{5}{4}. \tag{25.29}$$

La circulation de F le long de σ est donc égale à

$$\frac{4}{3} - \frac{5}{4} = \frac{1}{12}. \tag{25.30}$$

Comme prévu, nous obtenons le même résultat. △

25.2 Théorème de la divergence dans le plan

25.2.1 La convention de sens de parcours

Soient D, un domaine dans le plan et une paramétrisation

$$\begin{aligned} \sigma\colon [a,b] &\to \mathbb{R}^2 \\ t &\mapsto \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \end{aligned} \tag{25.31}$$

une paramétrisation du bord ∂D de D. La normale à σ est perpendiculaire à la tangente, donc la normale extérieure de norme 1 vaut

$$n = \frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}} \quad \text{ou} \quad n- = \frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}}. \tag{25.32}$$

Comment faire le choix ?

Nous prenons comme convention que le sens *du chemin* doit être tel que le vecteur normal extérieur soit

$$n = \frac{\big(y'(t), -x'(t)\big)}{\sqrt{\big(x'(t)\big)^2 + \big(y'(t)\big)^2}}. \tag{25.33}$$

Donc si le chemin σ 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 25.3.

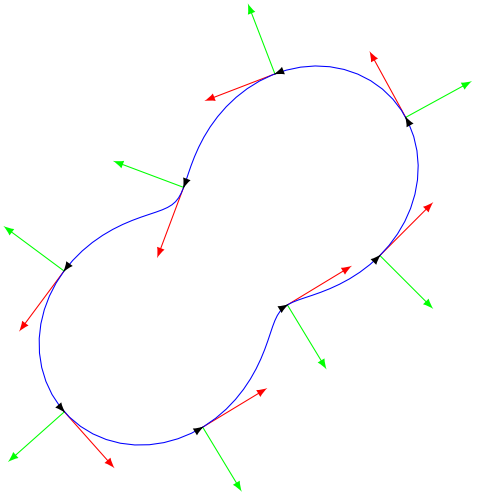


FIGURE 25.3 – Le champ de vecteurs tangents est dessiné en rouge tandis qu'en vert nous avons le champ de vecteurs normaux extérieurs.

25.2.2 Théorème de la divergence

Théorème 25.7 (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 :

$$\int_{\partial D} F \cdot n \, d\sigma = \int_D \nabla \cdot F \, dxdy. \tag{25.34}$$

Démonstration. Tant $F \cdot n$ que $\nabla \times F$ 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 :

$$\begin{aligned} F \cdot n &= \frac{1}{\|\sigma'\|} \begin{pmatrix} F_x \\ F_y \end{pmatrix} \cdot \begin{pmatrix} y' \\ -x' \end{pmatrix} \\ &= \frac{1}{\|\sigma'\|} (F_x y' - F_y x') \\ &= \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \\ &= \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'. \end{aligned} \tag{25.35}$$

Par conséquent, la *fonction*

$$F \cdot n \tag{25.36}$$

est la même que la *fonction*

$$\frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'. \tag{25.37}$$

L'intégrale de cette dernière fonction sur le chemin σ est

$$\begin{aligned} I &= \int_\sigma F \cdot n \\ &= \int_\sigma \frac{1}{\|\sigma'\|} \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma' \\ &= \int_a^b \frac{1}{\|\sigma'(t)\|} \begin{pmatrix} -F_y\big(\sigma(t)\big) \\ F_x\big(\sigma(t)\big) \end{pmatrix} \cdot \sigma'(t) \|\sigma'(t)\| dt \\ &= \int_a^b \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \cdot \sigma'(t) dt. \end{aligned} \tag{25.38}$$

Cette dernière intégrale est la circulation du champ de vecteurs $\begin{pmatrix} -F_y \\ F_x \end{pmatrix}$ sur le chemin σ. Le théorème de Green 25.5 nous enseigne que la circulation le long d'un chemin est égale au flux du rotationnel à travers la surface. Par conséquent,

$$I = \int_D \left(\nabla \times \begin{pmatrix} -F_y \\ F_x \end{pmatrix} \right) \cdot dS = \int_D \nabla \cdot F \, dxdy \tag{25.39}$$

□

25.3 Théorème de Stokes

Nous nous mettons maintenant dans $\mathbb{R}^3$, et nous y considérons une surface paramétrée S donc le bord est ∂S.

Théorème 25.8 (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 :

$$\int_S \nabla \times F \cdot dS = \int_{\partial S} F \cdot d\sigma. \tag{25.40}$$

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 25.5. Supposons maintenant que le bord reste plat, mais que la surface se déforme un petit peu. Le chemin

$$\sigma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 0 \end{pmatrix} \tag{25.41}$$

est tout autant le bord du disque plat de rayon 1 que celui de la demi-sphère

$$\phi(x, y) = \begin{pmatrix} x \\ y \\ \sqrt{1 - x^2 - y^2} \end{pmatrix}. \tag{25.42}$$

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 σ. 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 25.9

Soit C l'intersection entre le cylindre $x^2 + y^2 = 1$ et le plan $x + y + z = 1$. Calculer la circulation de

$$F(x, y, z) = \begin{pmatrix} -y^3 \\ x^3 \\ -z^3 \end{pmatrix} \tag{25.43}$$

le long de C.

Au lieu de calculer directement

$$\int_C F \cdot d\sigma, \tag{25.44}$$

nous allons calculer

$$\int_S \nabla \times F \cdot dS \tag{25.45}$$

où S est une surface dont C est le bord. Cette intégrale est à calculer avec la formule (21.278).

La première chose à faire est de trouver une surface dont le bord est C et en trouver une paramétrisation ϕ. Le plus simple est de prendre le graphe du plan sur le cercle $x^2 + y^2 + 1$. Une paramétrisation de cette surface est simplement

$$\begin{aligned} \phi\colon D &\to \mathbb{R}^3 \\ (x, y) &\mapsto \begin{pmatrix} x \\ y \\ 1 - x - y \end{pmatrix} \end{aligned} \tag{25.46}$$

où D est le disque de rayon 1. Étant donné que cela paramètre le plan $x + y + z - 1 = 0$, le vecteur normal est $n = e_x + e_y + z_z$. Nous pouvons cependant calculer ce vecteur normal en suivant la recette usuelle. D'abord les vecteurs tangents sont

$$\frac{\partial \phi}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \frac{\partial \phi}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{25.47}$$

Et le vecteur normal est donné par le produit vectoriel :

$$\begin{aligned} n &= \frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y} \\ &= \begin{vmatrix} e_x & e_y & e_z \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{vmatrix} \\ &= e_x + e_y + z_z. \end{aligned} \tag{25.48}$$

Ensuite, le rotationnel de F est donné par

$$\nabla \times F = 3(x^2 + y^2)e_z. \tag{25.49}$$

Par conséquent,

$$\nabla \times F \cdot \left(\frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y}\right) = 3(x^2 + y^2). \tag{25.50}$$

L'intégrale à calculer est donc

$$\begin{aligned} \int_S \nabla \times F \cdot dS &= \int_D (\nabla \times F)\big(\phi(x,y)\big) \cdot \left(\frac{\partial\phi}{\partial x} \times \frac{\partial\phi}{\partial y}\right) dxdy \\ &= 3\int_D (x^2 + y^2)dxdy. \end{aligned} \tag{25.51}$$

Cette dernière intégrale est l'intégrale d'une fonction sur le disque de rayon 1. Elle s'effectue en passant aux coordonnées polaires :

$$3\int_D (x^2 + y^2)dxdy = \int_0^{2\pi} d\theta \int_0^1 (r^2)r\,dr = \frac{3\pi}{2}. \tag{25.52}$$

△

25.4 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 25.10 (Théorème de la divergence ou de Gauss)**.**
Le flux d'un champ de vecteur F à travers une surface fermée est égale à l'intégrale de la divergence sur le volume correspondant :

$$\int_{\partial V} F \cdot dS = \int_V \nabla \cdot F\,dxdydz. \tag{25.53}$$

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 égal 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 25.11.
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.

25.12.
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

$$\int_{\partial V} uF \cdot n = \int V\nabla(uF) = \int_V u\nabla \cdot F + \int_V F \cdot \nabla u \tag{25.54}$$

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 :

$$\int_V F \cdot \nabla u = \int_{\partial V} uF \cdot n - \int_V u \nabla F. \tag{25.55}$$

Exemple 25.13
Calculer le flux du champ de vecteurs

$$F(x,y,z) = \begin{pmatrix} 2x \\ y^2 \\ z^2 \end{pmatrix} \tag{25.56}$$

à travers la sphère de rayon 1 centrée à l'origine. Nous utilisons le théorème de la divergence

$$\int_S F \cdot n \, dS = \int_B \nabla \cdot F \, dxdydz \tag{25.57}$$

où S est la sphère et B est la boule (la sphère pleine). La divergence de F se calcule :

$$\nabla \cdot F = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = 2 + 2x + 2y. \tag{25.58}$$

L'intégrale est donc en trois termes :

$$\begin{aligned} \int_B 2 = 2\text{Volume(B)} = \frac{8\pi}{3} \\ \int_B y \, dxdydz = 0 \\ \int_B z \, dxdydz = 0. \end{aligned} \tag{25.59}$$

△

Dans certains cas le théorème de Gauss permet de simplifier le calcul de l'intégrale d'une fonction sur une surface.

Exemple 25.14
Soit à calculer l'intégrale

$$I = \int_{\partial B} (x^2 + y + z) dS, \tag{25.60}$$

c'est-à-dire l'intégrale de la fonction $x^2 + y + z$ sur la sphère. Le vecteur normal à la sphère est

$$n = xe_x + ye_y + ze_z. \tag{25.61}$$

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

$$F = \begin{pmatrix} x \\ 1 \\ 1 \end{pmatrix}. \tag{25.62}$$

Nous pouvons donc simplement intégrer $\nabla \cdot F$ sur toute la boule :

$$I = \int_B \nabla \cdot F \, dxdydz = \int_B 1 \, dxdudz = \frac{4\pi}{3}. \tag{25.63}$$

△

25.5 Coordonnées curvilignes

25.5.1 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 général de ces systèmes de **coordonnées curvilignes**. Des coordonnées curvilignes sur $\mathbb{R}^3$ est n'importe quel [3] 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

$$x_i : \mathbb{R}^3 \to \mathbb{R}. \tag{25.64}$$

Les coordonnées cylindriques sont

$$\begin{cases} x_1(r,\theta,z) = r\cos\theta & \text{(25.65a)} \\ x_2(r,\theta,z) = r\sin\theta & \text{(25.65b)} \\ x_3(r,\theta,z) = z & \text{(25.65c)} \end{cases}$$

Soit donc un système général $q = (q_1, q_2, q_3)$ et

$$M(q) = \begin{pmatrix} x_1(q) \\ x_2(q) \\ x_3(q) \end{pmatrix}. \tag{25.66}$$

Si nous fixons q_2 et q_3 et que nous laissons varier q_1, nous obtenons une courbe [4] dont nous pouvons considérer le vecteur vitesse, c'est-à-dire le vecteur tangent. En chaque point nous avons ainsi trois vecteurs

$$\frac{\partial M}{\partial q_i}(q). \tag{25.67}$$

Nous disons que le système de coordonnées curviligne est **orthogonal** si ces trois vecteurs sont orthogonaux. Dans la suite nous supposerons que c'est toujours le cas.

Nous posons

$$h_i = \left\| \frac{\partial M}{\partial q_i} \right\| \tag{25.68}$$

et nous considérons les trois vecteurs normés

$$e_i = h_i^{-1} \frac{\partial M}{\partial q_i}. \tag{25.69}$$

Les trois vecteurs $\{e_1, e_2, e_3\}$ forment une base orthonormée dite **base locale**. Ce sont des vecteurs liés [5] au point M.

25.5.2 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

$$e_i \cdot e_j = \delta_{ij}. \tag{25.70}$$

Nous en étudions maintenant quelques conséquences. L'équation (25.69) peut s'écrire plus explicitement sous la forme

$$e_i = \sum_k h_i^{-1} \frac{\partial x_k}{\partial q_i} 1_k. \tag{25.71}$$

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 élément 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 :

$$A_{ki} = h_i^{-1}\frac{\partial x_k}{\partial q_i}. \tag{25.72}$$

Cela nous donne le changement de base

$$e_i = \sum_k A_{ki} 1_k. \tag{25.73}$$

Le produit $e_i \cdot e_j$ s'écrit alors

$$\begin{aligned} e_i \cdot e_j &= \sum_{kl} A_{ki}A_{lj} \underbrace{1_k \cdot 1_l}_{=\delta_{kl}} \\ &= \sum_{kl} A_{ki}A_{lj}\delta_{kl} \\ &= \sum_k A_{ki}A_{kj} \\ &= \sum_k (A^T)_{ik}A_{kj}. \end{aligned} \tag{25.74}$$

Or cela doit valoir δ_{ij}. Par conséquent

$$A^T = A^{-1}. \tag{25.75}$$

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 (25.73) en multipliant cette équation par $(A^{-1})_{il}$ des deux côtés et en faisant la somme sur i :

$$\sum_i (A^{-1})_{il} e_i = \sum_{kl} \underbrace{A_{ki}(A^{-1})_{il}}_{=\delta_{kl}} 1_k, \tag{25.76}$$

par conséquent

$$\sum_i (A^T)_{il} e_i = 1_l, \tag{25.77}$$

et

$$1_l = \sum_i A_{li} e_i = \sum_i h_i^{-1}\frac{\partial x_l}{\partial q_i} e_i. \tag{25.78}$$

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

$$\begin{aligned} e_1 \times e_2 &= e_3 \\ e_2 \times e_3 &= e_1 \\ e_3 \times e_1 &= e_2 \end{aligned} \tag{25.79}$$

Ces trois relations s'écrivent en une seule avec

$$e_i \times e_j = \sum_k \epsilon_{ijk} e_k \tag{25.80}$$

où

$$\epsilon_{ijk} = \begin{cases} 0 & \text{si } i,j,k \text{ ne sont pas tous différents} \\ 1 & \text{si } ijk \text{ se ramène à 123 par un nombre pair de permutations} \\ -1 & \text{si } ijk \text{ se ramène à 123 par un nombre impair de permutations} \end{cases} \tag{25.81}$$

est le **symbole de Levi-Civita**. La formule du produit vectoriel peut également être utilisée à l'envers sous la forme

$$e_k = \frac{1}{2}\sum_{ij} \epsilon_{ijk}\, e_i \times e_j. \tag{25.82}$$

Le symbole de Levi-Civita possède de nombreuses formules. En voici certaines, facilement démontrables en considérant tous les cas :

$$\epsilon_{ijk}\epsilon_{ijl} = \delta_{kl}|\epsilon_{ijk}|. \tag{25.83}$$

Grâce au symboles de Levi-Civita, le produit mixte des vecteurs de base a une belle forme :

$$e_l \cdot (e_i \times e_j) = \sum_k \epsilon_{ijk} e_l \times e_k = \sum_k \epsilon_{ijk}\delta_{lk} = \epsilon_{ijl}. \tag{25.84}$$

25.5.3 Coordonnées polaires

Les coordonnées curvilignes polaires sont données par

$$M(r,\theta) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}, \tag{25.85}$$

et par conséquent

$$\frac{\partial M}{\partial r} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}, \quad \frac{\partial M}{\partial \theta} = \begin{pmatrix} -r\sin(\theta) \\ r\cos(\theta) \end{pmatrix}. \tag{25.86}$$

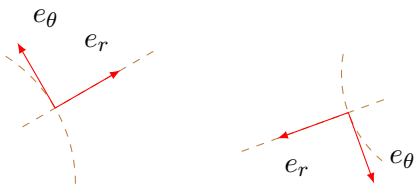
Nous avons les normes $h_r = 1$ et $h_\theta = r$, et donc les vecteurs de la base locale en (r,θ) sont

$$e_r = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} = \cos(\theta)e_x + \sin(\theta)e_y \tag{25.87}$$

ainsi que

$$e_\theta = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix} = -\sin(\theta)e_x + \cos(\theta)e_y. \tag{25.88}$$

Ces vecteurs sont représentés à la figure 25.4. Notez qu'il y en a une paire différente en chaque point.



(a) Base locale. (b) Base locale.

FIGURE 25.4 – 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_θ.

25.5.4 Coordonnées cylindriques

Les coordonnées cylindriques sont les mêmes que les coordonnées polaires à part qu'il faut écrire

$$M(r,\theta,z) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ z \end{pmatrix}, \tag{25.89}$$

et nous avons le vecteur de base supplémentaire

$$e_z = \frac{\partial M}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{25.90}$$

parce que $h_z = 1$.

25.5.5 Coordonnées sphériques

Les coordonnées curvilignes sphériques sont données par

$$M(\rho,\theta,\varphi)=\begin{pmatrix}\rho\sin(\theta)\cos(\varphi)\\ \rho\sin(\theta)\sin(\varphi)\\ \rho\cos(\theta)\end{pmatrix}, \tag{25.91}$$

dont les dérivées sont données par

$$\begin{aligned}\frac{\partial M}{\partial r}&=\begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix},\qquad \frac{\partial M}{\partial\theta}=\begin{pmatrix}\rho\cos(\theta)\cos(\varphi)\\ \rho\cos(\theta)\sin(\varphi)\\ -\rho\sin(\theta)\end{pmatrix},\\ \frac{\partial M}{\partial\varphi}&=\begin{pmatrix}-\rho\sin(\theta)\sin(\varphi)\\ \rho\sin(\theta)\cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{25.92}$$

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 (ρ,θ,φ) sont donc

$$\begin{aligned}e_r&=\begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix},\quad e_\theta=\begin{pmatrix}\cos(\theta)\cos(\varphi)\\ \cos(\theta)\sin(\varphi)\\ -\sin(\theta)\end{pmatrix},\\ e_\varphi&=\begin{pmatrix}-\sin(\varphi)\\ \cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{25.93}$$

25.5.6 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

$$\tilde{f}(q_1,q_2,q_3)=f\big(x_1(q),x_2(x),x_3(q)\big). \tag{25.94}$$

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 :

$$F(x,y,z)=\nabla f(x,y,z)=\begin{pmatrix}\partial_x f(x,y,z)\\ \partial_y f(x,y,z)\\ \partial_z f(x,y,z)\end{pmatrix}. \tag{25.95}$$

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

$$F(x,y,z)=F\big(x(q),y(q),z(q)\big)=\tilde{F}_1e_1+\tilde{F}_2e_2+\tilde{F}_3e_3. \tag{25.96}$$

Ces nombres seront des fonctions de (q_1,q_2,q_3).

Par définition,

$$\nabla f=\sum_l\frac{\partial f}{\partial x_l}1_l. \tag{25.97}$$

En remplaçant 1_l par sa valeur en termes des e_i par la formule (25.78),

$$\begin{aligned}\nabla f &= \sum_l \frac{\partial f}{\partial x_l} 1_l \\ &= \sum_l \frac{\partial f}{\partial x_l} \sum_i h_i^{-1} \frac{\partial x_l}{\partial q_i} e_i \\ &= \sum_{il} \frac{1}{h_i} \frac{\partial f}{\partial x_l} \frac{\partial x_l}{\partial q_i} e_i \\ &= \sum_i \frac{1}{h_i} \frac{\partial \tilde{f}}{\partial q_i} e_i.\end{aligned} \tag{25.98}$$

Plus explicitement,

$$\nabla f\big(x(q), y(q), z(q)\big) = \sum_i \frac{1}{h_i(q)} \frac{\partial \tilde{f}}{\partial q_i}(q) e_i \tag{25.99}$$

où

$$h_i(q) = \left\| \frac{\partial M}{\partial q_i}(q) \right\|. \tag{25.100}$$

Le plus souvent nous n'allons pas noter explicitement la dépendance de h_i en q.

25.5.6.1 Coordonnées sphériques

Nous pouvons exprimer le gradient d'une fonction en coordonnées sphériques en utilisant la formule (25.99) :

$$\nabla \tilde{f}(\rho, \theta, \varphi) = \frac{\partial \tilde{f}}{\partial \rho} e_\rho + \frac{1}{\rho} \frac{\partial \tilde{f}}{\partial \theta} e_\theta + \frac{1}{\rho \sin(\theta)} \frac{\partial \tilde{f}}{\partial \varphi} r_\varphi. \tag{25.101}$$

Cette expression peut paraitre peu pratique parce que les vecteurs e_ρ, e_θ et e_φ 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 25.15
Le potentiel de la gravitation est la fonction

$$V(x, y, z) = \frac{1}{\sqrt{x^2 + y^2 + z^2}}. \tag{25.102}$$

En coordonnées sphériques elle s'écrit

$$\tilde{V}(\rho, \theta, \varphi) = \frac{1}{\rho}. \tag{25.103}$$

En voilà une fonction qu'elle est facile à dériver, contrairement à V ! En suivant la formule (25.101), nous avons immédiatement

$$\nabla \tilde{V} = -\frac{1}{\rho^2} e_\rho. \tag{25.104}$$

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. △

25.5.7 Divergence en coordonnées curvilignes

Nous savons que

$$\nabla \tilde{f} = \sum_j \frac{1}{h_j} \frac{\partial \tilde{f}}{\partial q_j} e_j. \tag{25.105}$$

Nous pouvons en particulier considérer la fonction $f(q) = q_i$. De la même manière que nous avions 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

$$\nabla q_i = \sum_j \frac{1}{h_j} \frac{\partial q_i}{\partial q_j} e_j, \tag{25.106}$$

mais $\frac{\partial q_i}{\partial q_j} = \delta_{ij}$, donc

$$\nabla q_i = \frac{e_i}{h_i}, \tag{25.107}$$

ou encore

$$e_i = h_i \nabla q_i. \tag{25.108}$$

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 (25.82),

$$\nabla \cdot e_k = \frac{1}{2} \sum_{ij} \epsilon_{ijk} \nabla \cdot (e_i \times e_j). \tag{25.109}$$

Nous avons, en utilisant les règles de Leibnitz de la proposition 13.407,

$$\begin{aligned} \nabla \cdot (e_i \times e_j) &= \nabla \cdot (h_i \nabla q_i \times h_j \nabla q_j) \\ &= \nabla(h_i h_j) \cdot \big(\nabla q_i \times \nabla q_j\big) + h_i h_j \nabla \cdot \big(\nabla q_i \times \nabla q_j\big) \\ &= \nabla(h_i h_j) \cdot \big(\nabla q_i \times \nabla q_j\big) \\ &\quad + h_i h_j \nabla q_j \cdot \big(\underbrace{\nabla \times \nabla q_i}_{=0}\big) \\ &\quad + h_i h_j \nabla q_i \cdot \big(\underbrace{\nabla \times \nabla q_j}_{=0}\big) \end{aligned} \tag{25.110}$$

Cela nous fait

$$\nabla \cdot e_k = \sum_{ij} \epsilon_{ijk} \frac{\nabla(h_i h_j)}{h_i h_j} \cdot (e_i \times e_j). \tag{25.111}$$

parce que $\nabla q_i = h_i^{-1} e_i$. Nous pouvons développer le gradient qui intervient :

$$\nabla(h_i h_j) = \sum_l \frac{1}{h_l} \frac{\partial}{\partial q_l}(h_i h_j) e_l. \tag{25.112}$$

Nous voyons donc arriver le produit mixte $e_l \cdot (e_i \times e_j)$. En utilisant la formule (25.84), cela s'exprime directement sous la forme ϵ_{ijl}.

Nous avons alors

$$\begin{aligned} \nabla \cdot e_k &= \frac{1}{2} \sum_{ijl} \frac{1}{h_i h_j h_l} \frac{\partial}{\partial q_l}(h_i h_j) \epsilon_{ijk} \epsilon_{ijl} \\ &= \frac{1}{2} \sum_{ijl} \delta_{kl} |\epsilon_{ijk}| \frac{\partial}{\partial q_l}(h_i h_j) \\ &= \frac{1}{2} \sum_{ij} \frac{|\epsilon_{ijk}|}{h_i h_j h_k} \frac{\partial}{\partial q_k}(h_i h_j). \end{aligned} \tag{25.113}$$

Par exemple,

$$\nabla \cdot e_1 = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial q_1}(h_2 h_3). \tag{25.114}$$

Nous devons maintenant chercher le gradient d'un champ général

$$F(q) = \sum_k F_k(q) e_k. \tag{25.115}$$

La première chose à faire est d'utiliser la formule de Leibnitz :

$$\nabla \cdot F = \sum_k \nabla F_k(q) \cdot e_k + \sum_k F_k(q) \nabla \cdot e_k. \tag{25.116}$$

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 (25.113) signifie, c'est que

$$\nabla \cdot e_k = \frac{1}{h_1 h_2 h_3} \frac{\partial}{\partial q_k}(h_i h_j). \tag{25.117}$$

Nous savons déjà par la formule (25.99) que

$$\nabla F_k = \sum_l \frac{1}{h_l} \frac{\partial F_k}{\partial q_l} e_l, \tag{25.118}$$

par conséquent

$$\nabla F_k \cdot e_k = \sum_l \frac{1}{h_l} \frac{\partial F_k}{\partial q_l} \delta_{kl} = \frac{1}{h_k} \frac{\partial F_k}{\partial q_k}. \tag{25.119}$$

Pour obtenir cela nous avons utilisé le fait que $e_l \cdot e_k = \delta_{lk}$. Le terme numéro k de la somme (25.116) est donc

$$\frac{1}{h_k} \frac{\partial F_k}{\partial q_k} + \frac{F_k}{h_k h_i h_j} \frac{\partial (h_i h_j)}{\partial q_k} = \frac{1}{h_i h_j h_k} \frac{\partial (F_k h_i h_j)}{\partial q_k} \tag{25.120}$$

où il est entendu que i et j représentent les nombres tels que $\epsilon_{ijk} = 1$.

Au final, nous avons

$$\nabla \cdot F = \frac{1}{h_1 h_2 h_3} \sum_{ijk} |\epsilon_{ijk}| \frac{\partial (F_k h_i h_j)}{\partial q_k}. \tag{25.121}$$

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,

$$\nabla \cdot F = \frac{1}{h_1 h_2 h_3} \left[\frac{\partial}{\partial q_1}(F_1 h_2 h_3) + \frac{\partial}{\partial q_2}(F_2 h_1 h_3) + \frac{\partial}{\partial q_3}(F_3 h_1 h_2) \right]. \tag{25.122}$$

25.5.7.1 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

$$\nabla \cdot F = \frac{1}{r} \left[\frac{\partial}{\partial r}(r F_r) + \frac{\partial}{\partial \theta}(F_\theta) + \frac{\partial}{\partial z}(r F_z) \right]. \tag{25.123}$$

Par exemple si

$$F(r, \theta, z) = r e_\theta + e_z, \tag{25.124}$$

nous avons

$$(\nabla \cdot F)(r, \theta, z) = \frac{1}{r} \left[\frac{\partial}{\partial \theta}(r) + \frac{\partial}{\partial z}(r) \right] = 0. \tag{25.125}$$

Cela est logique parce que $r e_\theta$ est à peu près le champ dont nous avons parlé dans l'exemple (13.404), 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.

25.5.7.2 Coordonnées sphériques

En coordonnées sphériques, nous avons $h_\rho = 1$, $h_\theta = r$ et $h_\varphi = r \sin\theta$, donc

$$\nabla \cdot F = \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial \rho}(\rho^2 \sin\theta F_\rho) + \frac{\partial}{\partial \theta}(\rho \sin\theta F_\theta) + \frac{\partial}{\partial \varphi}(\rho F_\varphi) \right]. \tag{25.126}$$

si $F(\rho, \theta, \varphi) = F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi$.

25.5.8 Laplacien en coordonnées curvilignes orthogonales

Soit une fonction $f\colon \mathbb{R}^3 \to \mathbb{R}$. Le **laplacien** de f est donné par

$$\Delta f = \nabla \cdot (\nabla f). \tag{25.127}$$

En utilisant les formules données, nous avons

$$\Delta f = \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2h_3}{h_1}\frac{\partial f}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_1h_3}{h_2}\frac{\partial f}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1h_2}{h_3}\frac{\partial f}{\partial q_3}\right)\right]. \tag{25.128}$$

Dans cette expression, la fonction f est donnée comme fonction de q_1, q_2 et q_3.

En coordonnées cylindriques, cela s'écrit

$$\begin{aligned}\Delta f &= \frac{1}{r}\left[\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{\partial}{\partial \theta}\left(\frac{1}{r}\frac{\partial f}{\partial \theta}\right) + \frac{\partial}{\partial z}\left(r\frac{\partial f}{\partial z}\right)\right]\\ &= \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2}.\end{aligned} \tag{25.129}$$

Dans cette expression, f est fonction de r, θ et z.

En coordonnées sphériques, cela devient

$$\Delta f = \frac{1}{\rho^2\sin\theta}\left[\frac{\partial}{\partial \rho}\left(\rho^2\sin\theta\frac{\partial f}{\partial \rho}\right) + \frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial f}{\partial \theta}\right) + \frac{\partial}{\partial \varphi}\left(\frac{1}{\sin\theta}\frac{\partial f}{\partial \varphi}\right)\right]. \tag{25.130}$$

Dans cette expression, f est fonction de ρ, θ et φ.

25.5.9 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 13.407(3) avec F_kh_k en guise de f :

$$\begin{aligned}\nabla \times F_ke_k &= \nabla \times (F_kh_k\nabla q_k)\\ &= F_kh_k\underbrace{\nabla \times (\nabla q_k)}_{=0} + \nabla(F_kh_k) \times \nabla q_k\\ &= \frac{1}{h_k}\nabla(F_kh_k) \times e_k.\end{aligned} \tag{25.131}$$

Nous utilisons à présent la formule (25.99) du gradient et le formule $e_j \times e_k = \sum_l \epsilon_{jkl}e_l$:

$$\begin{aligned}\nabla \times (F_ke_k) &= \sum_j \frac{1}{h_jh_k}\frac{\partial}{\partial q_j}(F_kh_k)e_j \times e_k\\ &= \sum_{jl} \frac{1}{h_jh_k}\epsilon_{jkl}\frac{\partial}{\partial q_j}(F_kh_k)e_l.\end{aligned} \tag{25.132}$$

Le rotationnel s'écrit donc

$$\nabla \times F = \sum_{jkl} \frac{1}{h_jh_k}\epsilon_{jkl}\frac{\partial}{\partial q_j}(F_kh_k)e_l. \tag{25.133}$$

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

$$\frac{1}{h_2h_3}\left(\frac{\partial}{\partial q_2}(F_3h_3) - \frac{\partial}{\partial q_3}(F_2h_2)\right). \tag{25.134}$$

La formule complète devient

$$\begin{aligned}\nabla \times \sum_k F_ke_k = &\frac{1}{h_2h_3}\left(\frac{\partial}{\partial q_2}(F_3h_3) - \frac{\partial}{\partial q_3}(F_2h_2)\right)\\ &+ \frac{1}{h_1h_3}\left(\frac{\partial}{\partial q_3}(F_1h_1) - \frac{\partial}{\partial q_1}(F_3h_3)\right)\\ &+ \frac{1}{h_2h_1}\left(\frac{\partial}{\partial q_1}(F_2h_2) - \frac{\partial}{\partial q_2}(F_1h_1)\right).\end{aligned} \tag{25.135}$$

25.5.9.1 Coordonnées cylindriques

En utilisant $h_r = 1$, $h_\theta = r$ et $h_z = 1$, nous trouvons

$$\begin{aligned}\nabla \times (F_r e_r + F_\theta e_\theta + F_z e_z) = &\frac{1}{r}\left(\frac{\partial F_z}{\partial \theta} - \frac{\partial (F_\theta r)}{\partial z}\right) e_r \\ &+ \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right) e_\theta \\ &+ \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z.\end{aligned} \tag{25.136}$$

25.5.9.2 Coordonnées sphériques

En utilisant $h_\rho = 1$, $h_\theta = \rho$ et $h_\varphi = \rho \sin\theta$, nous trouvons

$$\begin{aligned}\nabla \times (F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi) = &\frac{1}{\rho \sin\theta}\left(\frac{\partial (F_\varphi)\sin\theta}{\partial \theta} - \frac{\partial F_\theta}{\partial \varphi}\right) e_\rho \\ &+ \frac{1}{\rho \sin\theta}\left(\frac{\partial F_\rho}{\partial \varphi} - \frac{\partial (F_\varphi \rho \sin\theta)}{\partial \rho}\right) e_\theta \\ &+ \frac{1}{\rho}\left(\frac{\partial F_\theta \rho}{\partial \rho} - \frac{\partial F_r}{\partial \theta}\right) e_\varphi.\end{aligned} \tag{25.137}$$

Note : dans le premier terme, il y a une simplification par ρ.

25.6 Les formules

25.6.1 Coordonnées polaires

Les vecteurs de base :

$$e_r = \begin{pmatrix}\cos(\theta) \\ \sin(\theta)\end{pmatrix} = \cos(\theta) e_x + \sin(\theta) e_y \tag{25.138a}$$

$$e_\theta = \begin{pmatrix}-\sin(\theta) \\ \cos(\theta)\end{pmatrix} = -\sin(\theta) e_x + \cos(\theta) e_y. \tag{25.138b}$$

Le gradient :

$$\nabla \tilde{f}(r,\theta) = \frac{\partial \tilde{f}}{\partial r}(r,\theta) e_r + \frac{1}{r}\frac{\partial \tilde{f}}{\partial \theta}(r,\theta) e_\theta. \tag{25.139}$$

La divergence :

$$\nabla \cdot F = \frac{1}{r}\left[\frac{\partial}{\partial r}(r F_r) + \frac{\partial}{\partial \theta}(F_\theta)\right]. \tag{25.140}$$

Le rotationnel :

$$\nabla \times (F_r e_r + F_\theta e_\theta) = \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z. \tag{25.141}$$

Notons que le rotationnel n'existe pas vraiment en deux dimensions. Ici nous avons vu le champ $F(r,\theta)$ comme un champs 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.

25.6.2 Coordonnées cylindriques

Les vecteurs de base : idem qu'en coordonnées polaires, et on ajoute e_z sans modifications.
Le gradient :

$$\nabla \tilde{f}(r,\theta,z) = \frac{\partial \tilde{f}}{\partial r}(r,\theta,z) e_r + \frac{1}{r}\frac{\partial \tilde{f}}{\partial \theta}(r,\theta,z) e_\theta + \frac{\partial \tilde{f}}{\partial z}(r,\theta,z) e_z. \tag{25.142}$$

La divergence :

$$\nabla \cdot F = \frac{1}{r}\left[\frac{\partial}{\partial r}(rF_r) + \frac{\partial}{\partial \theta}(F_\theta) + \frac{\partial}{\partial z}(rF_z)\right]. \tag{25.143}$$

Le rotationnel :

$$\begin{aligned}\nabla \times (F_r e_r + F_\theta e_\theta + F_z e_z) = &\frac{1}{r}\left(\frac{\partial F_z}{\partial \theta} - \frac{\partial (F_\theta r)}{\partial z}\right) e_r \\ &+ \left(\frac{\partial F_r}{\partial z} - \frac{\partial F_z}{\partial r}\right) e_\theta \\ &+ \left(\frac{\partial (F_\theta r)}{\partial r} - \frac{\partial F_r}{\partial \theta}\right) e_z.\end{aligned} \tag{25.144}$$

Note : les formules concernant les coordonnées polaires se réduisent de celles-ci en enlevant toutes les références à z.

25.6.3 Coordonnées sphériques

Les vecteurs de base :

$$\begin{aligned} e_r &= \begin{pmatrix}\sin(\theta)\cos(\varphi)\\ \sin(\theta)\sin(\varphi)\\ \cos(\theta)\end{pmatrix}, \quad e_\theta = \begin{pmatrix}\cos(\theta)\cos(\varphi)\\ \cos(\theta)\sin(\varphi)\\ -\sin(\theta)\end{pmatrix}, \\ e_\varphi &= \begin{pmatrix}-\sin(\varphi)\\ \cos(\varphi)\\ 0\end{pmatrix}\end{aligned} \tag{25.145}$$

Le gradient :

$$\nabla \tilde{f}(\rho, \theta, \varphi) = \frac{\partial \tilde{f}}{\partial \rho} e_\rho + \frac{1}{\rho}\frac{\partial \tilde{f}}{\partial \theta} e_\theta + \frac{1}{\rho \sin(\theta)}\frac{\partial \tilde{f}}{\partial \varphi} r_\varphi. \tag{25.146}$$

La divergence :

$$\nabla \cdot F = \frac{1}{\rho^2 \sin\theta}\left[\frac{\partial}{\partial \rho}(\rho^2 \sin\theta F_\rho) + \frac{\partial}{\partial \theta}(\rho \sin\theta F_\theta) + \frac{\partial}{\partial \varphi}(\rho F_\varphi)\right]. \tag{25.147}$$

Le rotationnel :

$$\begin{aligned}\nabla \times (F_\rho e_\rho + F_\theta e_\theta + F_\varphi e_\varphi) = &\frac{1}{\rho \sin\theta}\left(\frac{\partial (F_\varphi)\sin\theta}{\partial \theta} - \frac{\partial F_\theta}{\partial \varphi}\right) e_\rho \\ &+ \frac{1}{\rho \sin\theta}\left(\frac{\partial F_\rho}{\partial \varphi} - \frac{\partial (F_\varphi \rho \sin\theta)}{\partial \rho}\right) e_\theta \\ &+ \frac{1}{\rho}\left(\frac{\partial F_\theta \rho}{\partial \rho} - \frac{\partial F_r}{\partial \theta}\right) e_\varphi.\end{aligned} \tag{25.148}$$

Chapitre 26

Espaces de Hilbert

26.1 Espaces de Hilbert

Définition 26.1.
Un espace ***préhilbertien*** *est*

— *soit un espace vectoriel réel muni d'un produit scalaire*[1],

— *soit un espace vectoriel complexe muni d'un produit scalaire hermitien*[2].

Un ***espace de Hilbert*** *est un espace préhilbertien qui est complet*[3] *pour la norme induite par son produit (scalaire ou hermitien), c'est-à-dire tel que*

$$\|x\| = \sqrt{\langle x, x\rangle} \tag{26.1}$$

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 9.43.

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.

Proposition 26.2.
Si $\mathscr{H}$ est un espace de Hilbert réel, alors

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle. \tag{26.2}$$

Si $\mathscr{H}$ est un espace de Hilbert complexe alors

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\langle x, y\rangle. \tag{26.3}$$

Dans les deux cas nous avons l'inégalité de Cauchy-Schwarz :

$$|\langle x, y\rangle| \leqslant \|x\|\|y\|. \tag{26.4}$$

Dans un espace vectoriel de dimension infinie, tous les opérateurs linéaires ne sont pas continus.

Exemple 26.3
Soit un espace vectoriel V engendré par la base $\{e_k\}_{k\in\mathbb{N}}$ et l'application linéaire $T\colon V \to V$ donnée par

$$Te_k = ke_k. \tag{26.5}$$

1. Définition 11.5.
2. Définition 11.8.
3. Définition 9.20.

Nous allons montrer que l'image inverse de la boule unité ouverte $\mathcal{O}$ n'est pas ouverte. En effet si (ϵ_k) est une suite de réels strictement positifs tendant vers zéro, les vecteurs

$$a_k = \left(\frac{1}{k} + \epsilon_k\right) e_k \tag{26.6}$$

sont hors de $T^{-1}\mathcal{O}$ parce que

$$Ta_k = (1 + k\epsilon_k)e_k. \tag{26.7}$$

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. △

Proposition 26.4 ([1])**.**
Soit un espace de Hilbert $\mathscr{H}$ et un espace vectoriel normé $(E, \|.\|)$ tel qu'il existe une bijection linéaire isométrique $\Phi\colon \mathscr{H} \to E$. Alors en posant

$$\langle \alpha, \beta \rangle_E = \langle \Phi^{-1}(\alpha), \Phi^{-1}(\beta) \rangle_{\mathscr{H}}, \tag{26.8}$$

l'espace E devient un espace de Hilbert.

Démonstration. Le fait que $\langle ., . \rangle_E$ soit un produit scalaire (hermitien) lorsque $\langle ., . \rangle_{\mathscr{H}}$ l'est est une conséquence de la linéarité de Φ. Il donne la norme parce que

$$\sqrt{\langle \alpha, \alpha \rangle} = \sqrt{\langle \Phi^{-1}(\alpha), \Phi^{-1}(\alpha) \rangle} = \|\Phi^{-1}(\alpha)\| = \|\alpha\| \tag{26.9}$$

parce que Φ^{-1} est une isométrie.

Pour voir que E est complet, soit une suite (α_k) de τ-Cauchy dans E. Nous devons prouver qu'elle converge. Nous posons $x_k = \Phi^{-1}(\alpha_k)$.

Commençons par voir que (x_k) est τ-Cauchy dans $\mathscr{H}$. Soit un voisinage V de 0 dans $\mathscr{H}$. Nous posons $U = \Phi(V)$. Vu que (α_k) est τ-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 $\Phi(\alpha_k - \alpha_l) \in \Phi(U) = V$.

Vu que $\mathscr{H}$ est complet, il existe $x \in \mathscr{H}$ tel que $x_k \xrightarrow{\mathscr{H}} x$. Montrons que $\alpha_k \xrightarrow{\mathscr{H}'} \Phi(x)$.

Soit $\epsilon > 0$ et $N > 0$ tel que $\|x_k - x\| < \epsilon$ dès que $k \geqslant N$. Alors nous avons aussi

$$\|\alpha_k - \Phi(x)\| = \|\Phi(x_k) - \Phi(x)\| = \|\Phi(x_k - x)\| = \|x_k - x\| < \epsilon \tag{26.10}$$

où nous avons utilisé le fait que Φ était linéaire et une isométrie. □

26.1.1 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, 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 $\mathscr{H}$. L'ensemble $\mathrm{Span}(A)$ des combinaisons linéaires d'éléments de A est un sous-espace vectoriel de $\mathscr{H}$, 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

$$a = \sum_{k=1}^{\infty} \alpha_k v_k \tag{26.11}$$

où les α_k sont suffisamment décroissants pour assurer les convergences [4]. Le vecteur a n'est pas dans $\mathrm{Span}(A)$, mais la suite $a_n = \sum_{k=1}^{n} \alpha_k v_k$ est dans $\mathrm{Span}(A)$ et converge vers a : $a_n \xrightarrow{\mathscr{H}} a$. En effet,

$$\|a_n - a\| = \|\sum_{k=n+1}^{\infty} \alpha_k v_k\| \leqslant \sum_{k=n+1}^{\infty} |\alpha_k| \xrightarrow{n\to\infty} 0. \tag{26.12}$$

4. Par exemple $\alpha_k = 1/k^2$ si les v_k sont orthonormaux.

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 [5] converge vers zéro :

$$\lim_{n\to\infty} \sum_{k=n+1}^{\infty} |\alpha_k| = 0. \tag{26.13}$$

26.2 Théorème de la projection

Théorème 26.5 (Projection sur partie fermée convexe[385, 386])**.**
Soit $\mathscr{H}$ un espace de Hilbert, $x \in \mathscr{H}$, et C un sous-ensemble fermé convexe de $\mathscr{H}$.

(1) Les deux conditions suivantes sur $y \in \mathscr{H}$ sont équivalentes :

(1a) $\|x - y\| = \inf\{\|x - z\| \text{ tel que } z \in C\}$,

(1b) pour tout $z \in C$, $\mathrm{Re}\langle x - y, z - y\rangle \leqslant 0$.

(2) Il existe un unique $y \in \mathscr{H}$, noté $y = \texttt{proj}_C(x)$ 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

$$\lim_{n\to\infty} \|x - y_n\| = \inf\{\|x - y\| \text{ tel que } z \in C\} = d. \tag{26.14}$$

Nous allons montrer que cette suite peut être choisie de Cauchy. Elle convergera donc dans $\mathscr{H}$ parce que ce dernier est complet. Mais C étant supposé fermé dans $\mathscr{H}$, la limite appartiendra à C. Soient $r, s \in \mathbb{N}$. D'abord nous avons

$$\begin{aligned}\|y_r - y_s\|^2 &= \langle y_r - y_s + x - x, y_r - y_s + x - x\rangle &\text{(26.15a)}\\ &= \|y_r - x\|^2 + \|y_s - x\|^2 - 2\langle y_r - x, y_s - x\rangle. &\text{(26.15b)}\end{aligned}$$

Ensuite,

$$\begin{aligned}4\left\|\frac{y_r + y_s}{2} - x\right\|^2 &= \langle y_r + y_s - 2x, y_r - y_s - 2x\rangle &\text{(26.16a)}\\ &= \|y_r - x\|^2 + \|y_s - x\|^2 + 2\langle y_r - x, y_s - x\rangle. &\text{(26.16b)}\end{aligned}$$

Si nous égalisons les valeurs de $2\langle y_r - x, y_s - x\rangle$ nous trouvons

$$\|y_r - y_s\|^2 = -4\left\|\frac{y_r + y_s}{2} - x\right\|^2 + 2\|y_r - x\|^2 + 2\|y_s - x\|^2. \tag{26.17}$$

La distance infimum étant d, nous pouvons choisir y_n de telle façon à avoir

$$\|y_n - x\| \leqslant d + \frac{1}{n}. \tag{26.18}$$

D'autre part étant donné que C est convexe, $(y_r + y_s)/2$ est dans C et nous avons

$$\left\|\frac{y_r + y_s}{2} - x\right\| \leqslant d. \tag{26.19}$$

En mettant ces majorations dans (26.17) nous trouvons

$$\|y_r - y_s\|^2 \leqslant -4d + 2\left(d + \frac{1}{r}\right) + 2\left(d + \frac{1}{s}\right) = \frac{1}{r} + \frac{1}{s}. \tag{26.20}$$

5. On se comprend hein.

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

$$d \leqslant \|x-y\| \leqslant \|x-y_n\| + \|y_n - y\|. \tag{26.21}$$

En prenant le limite $n \to \infty$ nous trouvons

$$d \leqslant \|x-y\| \leqslant d. \tag{26.22}$$

Unicité Même preuve que pour le théorème en dimension finie 13.233.

(1a)⇒ (1b) Même preuve que pour le théorème en dimension finie 13.233.

(1b)⇒ (1a) Même preuve que pour le théorème en dimension finie 13.233.

□

Proposition 26.6.
Soit C une partie convexe et fermée de l'espace de Hilbert $\mathscr{H}$. Alors pour tout $x_1, x_2 \in \mathscr{H}$ nous avons

$$\| \texttt{proj}_C(x_1) - \texttt{proj}_C(x_2)\| \leqslant \|x_1 - x_2\|. \tag{26.23}$$

En particulier la projection est une application continue.

Démonstration. Nous posons $y_1 = \texttt{proj}_C(x_1)$ et $y_2 = \texttt{proj}_C(x_2)$. Par la partie (1b) du théorème 26.5, nous avons, pour tout $z, z' \in C$ les inégalités

$$\text{Re}\langle x_1 - y_1, z - y_1\rangle \leqslant 0 \tag{26.24a}$$
$$\text{Re}\langle x_2 - y_2, z' - y_2\rangle \leqslant 0. \tag{26.24b}$$

En prenant $z = y_2$ et $z' = y_1$ et en sommant nous trouvons

$$\text{Re}\langle (x_1 - y_1) + (y_2 - x_2), y_2 - y_1\rangle \leqslant 0. \tag{26.25}$$

Nous pouvons maintenant calculer

$$\begin{aligned}
\|y_1 - y_2\|^2 &= \text{Re}\, \|y_1 - y_2\| \\
&= \text{Re}\langle y_1 - y_2, (y_1 - x_1) + x_1 - x_2 + (x_2 - y_2)\rangle \\
&= \text{Re}\langle y_1 - y_2, x_1 - x_2\rangle + \text{Re}\langle y_2 - y_1, (x_1 - y_1) + (y_2 - x_2)\rangle \\
&\leqslant \langle y_1 - y_2, x_1 - x_2\rangle \\
&\leqslant \|\langle y_1 - y_2, x_1 - x_2\rangle\| \\
&=\leqslant \|y_1 - y_2\|\|x_1 - x_2\|.
\end{aligned} \tag{26.26}$$

En simplifiant par $\|y_1 - y_2\|$ [6] nous trouvons le résultat

$$\|y_1 - y_2\| \leqslant \|x_1 - x_2\|. \tag{26.27}$$

□

Théorème 26.7 (Projection orthogonale)**.**
Soit $\mathscr{H}$ un espace de Hilbert et K, un sous-espace vectoriel fermé non réduit à $\{0\}$ et $x \in \mathscr{H}$. L'élément $y = \texttt{proj}_K(x)$ est l'unique élément de K tel que

$$x - y \in K^\perp. \tag{26.28}$$

De plus l'application $x \mapsto \texttt{proj}_K(x)$ est linéaire, continue et de norme 1.

L'élément y ainsi définit est la **projection orthogonale** de x sur K et sera noté $\texttt{proj}_K(f)$.

6. Si c'est nul, alors la preuve est évidente.

Démonstration. La continuité de la projection est donné par la proposition 26.6.

Soit z un élément de K tel que $\langle z-x, a\rangle = 0$ pour tout $a \in K$. Nous avons

$$\|x-a\|^2 = \|z-x\|^2 + \|a-z\|^2 + 2\underbrace{\langle z-x, a-z\rangle}_{=0} \tag{26.29a}$$

$$\geqslant \|z-x\|^2. \tag{26.29b}$$

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

$$\|(y+\lambda a) - x\|^2 - \|y-x\|^2 \tag{26.30}$$

est positive. En développant les produits scalaires nous trouvons la conditions suivante

$$\lambda^2\|a\|^2 + 2\lambda\langle a, y-x\rangle \geqslant 0 \tag{26.31}$$

qui doit être vraie pour tout $\lambda \in \mathbb{R}$. En tant que polynôme du second degré en λ, 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 $x_1, x_2 \in \mathscr{H}$. L'élément $y = \mathtt{proj}_K x_1 + \mathtt{proj}_K x_2$ satisfait à la condition d'orthogonalité : pour tout $z \in K$,

$$\langle x_1 + x_2 - \mathtt{proj}_K x_1 - \mathtt{proj}_K x_2, z\rangle = \langle x_1 - \mathtt{proj}_K x_1, z\rangle + \langle x_2 - \mathtt{proj}_K x_2, z\rangle = 0. \tag{26.32}$$

Étant donné que K est un sous-espace vectoriel, la condition de minimalité est automatiquement vérifiée (seconde partie du théorème 26.5).

En ce qui concerne la norme opérateur de $\mathtt{proj}_K$, la décomposition de $x \in \mathscr{H}$ en composantes dans K et $K^\perp$ est

$$x = x + (x - \mathtt{proj}_K x). \tag{26.33}$$

Étant deux parties orthogonales nous avons

$$\|\mathtt{proj}_K x\|^2 = \|x\|^2 - \|x - \mathtt{proj}_K x\|^2. \tag{26.34}$$

En prenant $\|x\| = 1$ nous trouvons $\|\mathtt{proj}_K x\|^2 \leqslant 1$ et par conséquent $\|\mathtt{proj}_K\| \leqslant 1$. Mais d'autre part en prenant $x \in K$ nous avons automatiquement $\|\mathtt{proj}_K\| \geqslant 1$. □

Proposition 26.8.
Soit $\mathscr{H} = L^2(\Omega, \mathcal{A}, \mu)$, $\mathcal{F}$ 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 $\mathtt{proj}_K f$ est positive (presque partout).

Démonstration. L'ensemble $A = \{\mathtt{proj}_K f < 0\}$ est dans $\mathcal{F}$. En effet

$$A = (\mathtt{proj}_K f)^{-1}\big(]-\infty, 0[\big) \tag{26.35}$$

alors que, par construction, $\mathtt{proj}_K f$ est $\mathcal{F}$-mesurable. La fonction indicatrice $\mathbb{1}_A$ est alors $\mathcal{F}$-mesurable (c'est-à-dire $1_A \in K$) et nous avons

$$0 \leqslant \int_\Omega f\mathbb{1}_A = \int_\Omega \mathtt{proj}_K f\mathbb{1}_A \leqslant 0. \tag{26.36}$$

Étant donné que nous avons supposé $f \geqslant 0$ nous avons alors $\mu(A) = 0$. D'où le fait que $\mathtt{proj}_K f$ est presque partout positive. □

26.3 Systèmes orthogonaux et bases

Dans cette partie nous noterons $\mathbb{K}$ le corps de base de l'espace $\mathscr{H}$. Seuls deux cas sont envisagés : $\mathbb{K} = \mathbb{C}$ ou $\mathbb{K} = \mathbb{R}$.

Pour chaque $x \in \mathscr{H}$ nous considérons l'application

$$\begin{aligned} \Phi_y : \mathscr{H} &\to \mathbb{K} \\ x &\mapsto \langle x, y \rangle. \end{aligned} \tag{26.37}$$

Ce sont des applications continues.

26.3.1 Orthogonal d'une partie

Définition 26.9.
*Soit une partie A de l'espace de Hilbert $\mathscr{H}$. L'**orthogonal** de A est l'ensemble*

$$A^\perp = \{v \in \mathscr{H} \text{ tel que } \langle v, x \rangle = 0 \forall x \in A\}. \tag{26.38}$$

Proposition 26.10.
Si $\mathscr{H}$ est une préhilbert et si $A \subset \mathscr{H}$, alors l'ensemble $A^\perp$ est un sous-espace fermé de $\mathscr{H}$.

Démonstration. L'application Φ_x définie en (26.37) est continue pour chaque $x \in \mathscr{H}$ et par conséquent l'ensemble $\ker \Phi_x = \Phi_x^{-1}(\{0\})$ est fermé. L'ensemble

$$A^\perp = \bigcap_{x \in A} \ker \Phi_x \tag{26.39}$$

est donc fermé. □

Proposition 26.11 (wikipédia)**.**
Soit V un espace vectoriel normé et une décomposition en somme directe $V = F \oplus G$. Alors les trois conditions suivantes sont équivalentes.

(1) L'isomorphisme naturel $F \times G \to v$ est un homéomorphisme.

(2) F et G sont fermés et la restriction à G de la projection $V \to V/F$ est un homéomorphisme.

(3) F et G sont fermés et la projection de E sur F est continue.

Lorsqu'une décomposition en somme directe vérifie la proposition 26.11, nous disons que la décomposition est **topologique**.

Théorème 26.12.
Soit $\mathscr{H}$ un espace de Hilbert et F un sous-espace fermé de $\mathscr{H}$. Alors

(1) Nous avons la décomposition

$$\mathscr{H} = F \oplus F^\perp. \tag{26.40}$$

(2) La projection sur F par rapport à la somme directe (26.40) est la projection $\texttt{proj}_F$ du théorème de projection.

(3) La décomposition (26.40) est topologique.

Proposition 26.13.
Si F est un sous-espace de l'espace de Hilbert $\mathscr{H}$ alors on a $F^{\perp\perp} = \bar{F}$.

Démonstration. Nous savons par la proposition 26.10 que $F^\perp$ est fermé, par conséquent le théorème 26.12 donne la somme directe

$$\mathscr{H} = F^\perp \oplus F^{\perp\perp}. \tag{26.41}$$

Mais $\bar{F}$ étant également fermé nous avons la somme directe

$$\mathscr{H} = \bar{F} \oplus (\bar{F})^{\perp}. \tag{26.42}$$

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 dons $\langle x, y \rangle = 0$ par continuité du produit scalaire.

Nous avons donc

$$\mathscr{H} = F^{\perp} \oplus F^{\perp\perp} = \bar{F} \oplus F^{\perp}. \tag{26.43}$$

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. □

Proposition 26.14.
Soit $\mathscr{H}$ un espace de Hilbert et F un sous-espace vectoriel de $\mathscr{H}$. Alors F est dense si et seulement si $F^{\perp} = \{0\}$.

Démonstration. Nous savons que

$$\mathscr{H} = F^{\perp} \oplus \bar{F}. \tag{26.44}$$

Donc nous avons $H = \bar{F}$ si et seulement si $F^{\perp} = \{0\}$. □

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.

26.3.2 Dual, théorème de représentation de Riesz

Lemme 26.15.
L'application linéaire

$$\begin{aligned} \Phi\colon \mathscr{H} &\to \mathcal{L}(\mathscr{H}, \mathbb{K}) \\ y &\mapsto \Phi_y \end{aligned} \tag{26.45}$$

est une isométrie : nous avons $\|\Phi_y\| = \|y\|$.

De plus pour chaque y, l'application Φ_y est continue.

Démonstration. En utilisant la définition de la norme opérateur et l'inégalité de Cauchy-Schwarz,

$$\|\Phi_y\| = \sup_{\|x\|=1} \|\Phi_y(x)\| = \sup |\langle x, y \rangle| \leqslant \sup \|x\|\|y\| = \|y\|. \tag{26.46}$$

Par conséquent $\|\Phi_y\| \leqslant \|y\|$. Mais d'autre part le fait que $\Phi_y(y) = \|y\|^2$ montre que $\|\Phi_y\| \geqslant \|y\|$.

En ce qui concerne la continuité de Φ_y, elle est garantie par le fait que c'est une application linéaire bornée via la proposition 12.25. □

Définition 26.16.
*Le **dual** de l'espace de Hilbert $\mathscr{H}$ est l'ensemble*

$$\mathscr{H}' = \{f\colon \mathscr{H} \to \mathbb{K} \text{ } \textit{linéaire et continue}\}. \tag{26.47}$$

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.

Théorème 26.17 (Théorème de représentation de Riesz, thème 33)**.**
Soit un espace de Hilbert $\mathscr{H}$ sur le corps $\mathbb{K}$ ($\mathbb{R}$ ou $\mathbb{C}$). L'application

$$\begin{aligned}\Phi\colon \mathscr{H} &\to \mathscr{H}'\\ y &\mapsto \Phi_y\end{aligned} \tag{26.48}$$

est une bijection isométrique.

Démonstration. Nous savons du lemme 26.15 que Φ est une isométrie. Nous devons seulement montrer que Φ est surjective. L'application nulle est dans l'image de Φ. Soit $f \in \mathscr{H}'$ non nulle. Par continuité nous savons que $F = \ker(f)$ est fermé, donc

$$\mathscr{H} = \ker(f) \oplus (\ker f)^\perp \tag{26.49}$$

par le théorème 26.12.

Une base adaptée Nous considérons une base de $\mathscr{H}$ adaptée à la décomposition 26.49, c'est-à-dire

- $(u_s)_{s\in S}$ une base de $\ker f$,
- $(v_t)_{t\in T}$ une base de $(\ker f)^\perp$.

$(\ker f)^\perp \cap \ker(f) = \{0\}$ 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 $v = 0$.

Une base encore plus adaptée Si v est non nul dans $\in (\ker f)^\perp$, 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 4.39) assure alors que

$$\{v\} \cup \{u_s\}_{s\in S} \tag{26.50}$$

est une base de $\mathscr{H}$ avec $v \in (\ker f)^\perp$ et $u_s \in \ker f$. Nous choisissons v pour avoir $\|v\| = 1$.

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 :

$$\Phi_y(v) = \langle v, y\rangle = f(v)\langle v, v\rangle = f(v) \tag{26.51a}$$
$$\Phi_y(u_s) = \langle u_s, y\rangle = 0. \tag{26.51b}$$

Par conséquent Φ_y et f coïncident sur une base de $\mathscr{H}$.

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

$$\Phi_y(v) = \bar\lambda\langle v, v\rangle = \bar\lambda. \tag{26.52}$$

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 $\mathbb{K}^n$.

26.3.3 Séparabilité

Définition 26.18.
Un espace topologique est ***séparable*** *s'il possède une partie dénombrable dense.*

Définition 26.19 ([387]).
Si E est un espace vectoriel normé nous disons que Δ est une partie ***totale*** *de E si* $\mathrm{Span}(\Delta)$ *est dense dans E. Attention : nous rappelons que* $\mathrm{Span}(\Delta)$ *est l'ensemble des combinaisons linéaires* finies *d'éléments de Δ.*

Proposition 26.20.
Un espace vectoriel normé est séparable si et seulement s'il possède une partie totale dénombrable.

Démonstration. Si Δ est une partie dénombrable de E, alors le $\mathbb{Q}$-espace vectoriel $\mathrm{Span}_{\mathbb{Q}}\Delta$ est dénombrable, et sa fermeture est la même que celle de $\mathrm{Span}_{\mathbb{K}}\Delta$. □

Définition 26.21.
Une famille $(u_i)_{i\in I}$ d'éléments de $\mathscr{H}$ indicée 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 26.2, nous avons

$$\Big\| \sum_{k=1}^{n} \xi_k u_k \Big\|^2 = \sum_{k=1}^{n} |\xi_k|^2. \tag{26.53}$$

Exemple 26.22
Dans l'ensemble $L^2(a,b)$ avec $b-a=L$ l'ensemble des fonctions

$$e_n(t) = \frac{1}{\sqrt{b-a}} \exp\Big(\frac{2\pi}{L} int\Big) \tag{26.54}$$

avec $n \in \mathbb{Z}$ forme un système orthonormé. En effet

$$\langle e_n, e_n\rangle = \frac{1}{b-a}\int_a^b \exp\Big(\frac{2\pi}{L} int\Big)\exp\Big(-\frac{2\pi}{L} int\Big) = 1 \tag{26.55}$$

et

$$\langle e_n, e_m\rangle = \frac{1}{b-a}\int_a^b e^{\frac{2\pi}{L}(n-m)t} dt = 0. \tag{26.56}$$

Dans cette intégrale nous utilisons le fait que $b = a + (b-a)$ pour simplifier les expressions en cours de calcul.

La famille (26.54) est le **système trigonométrique** de $L^2(a,b)$. On en parle aussi dans [351]. △

Proposition 26.23.
Une partie orthonormée est libre.

Démonstration. Soit $(u_i)_{i\in I}$ une famille orthonormée de $\mathscr{H}$ et une combinaison linéaire *finie* nulle :

$$\sum_{k=1}^{n} a_k u_k = 0. \tag{26.57}$$

Nous développons la somme en utilisant les formules de la proposition 26.2 :

$$0 = \Big\| \sum_k a_k u_k \Big\|^2 = \sum_k |a_k|^2 \|u_k\|^2 + 2\sum_{k<l} \langle a_k u_k, a_l u_l\rangle. \tag{26.58}$$

La famille étant orthonormée, les choses se simplifient en

$$\sum_k |a_k|^2 = 0, \tag{26.59}$$

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

$$\sum_{i\in I} a_i = \sup_{\substack{J\subset I\\ J \text{ finie}}} \sum_{j\in J} a_j. \tag{26.60}$$

Cela est discuté dans la section 12.7.

Proposition 26.24 (Inégalités de Bessel)**.**
Soit $\mathscr{H}$ un préhilbert. Si $(u_i)_{i\in I}$ est un système orthonormé et si $x \in \mathscr{H}$, alors

$$\sum_{i\in I} \big|\langle x, u_i\rangle\big|^2 \leqslant \|x\|^2. \tag{26.61}$$

Démonstration. Les éléments de la somme étant des réels positifs, la notion de somme à utiliser est celle de la définition 12.99.

Posons $c_i(x) = \langle x, u_i\rangle$. Pour toute partie finie $J \subset I$ nous avons

$$0 \leqslant \Big\|x - \sum_{j\in J} c_j(x)u_j\Big\|^2 = \|x\|^2 - 2\operatorname{Re}\sum_j \langle x, c_j(x)u_j\rangle + \sum_j |c_j(x)|^2. \tag{26.62}$$

Mais en tenant compte du fait que

$$\langle x, c_j(x)u_j\rangle = \overline{c_j(x)}\langle x, u_j\rangle = |c_j(x)|^2, \tag{26.63}$$

nous restons avec

$$\|x - \sum_{j\in J} c_j(x)u_j\| = \|x\|^2 - \sum_{j\in J} |c_j(x)|^2. \tag{26.64}$$

Finalement,

$$\sum_{j\in J} |c_j(x)|^2 \leqslant \|x\|^2. \tag{26.65}$$

Ayant cette inégalité pour toute partie finie de I, nous l'avons encore pour le supremum. □

Proposition 26.25.
Soit $\mathscr{H}$ un préhilbert et une famille orthonormé $(u_i)_{i\in I}$. Si

$$x = \sum_{i\in I} \xi_i u_i \tag{26.66}$$

alors $\xi_i = \langle x, u_i\rangle$.

Démonstration. Nous appliquons l'application Φ_{u_k} du théorème de représentation de Riesz[7] à l'équation (26.66).

Si I est dénombrable, alors permuter Φ_{u_k} avec la somme consiste à invoquer la continuité, et permuter la limite des sommes partielles avec Φ_{u_k} (l'application Φ_{u_k} est continue parce qu'isométrique).

Sinon, il faut utiliser la proposition 12.108. Il faut donc montrer que la famille $\Phi_{u_k}\big(\xi_i u_i\big)$ est sommable. Cela est fort vrai parce que cette famille ne contient en réalité qu'en seul élément non nul, celui avec $i = k$, qui vaut ξ_k. Au final nous avons :

$$\langle x, u_k\rangle = \sum_{i\in I} \Phi_{u_k}(\xi_i u_i) = \xi_k. \tag{26.67}$$

□

7. Théorème 26.17.

26.3.4 Base hilbertienne

Définition 26.26.
Une ***base orthonormée*** *est une famille dénombrable orthonormé et totale*[8]*. Cela sera souvent aussi appelé une* ***base hilbertienne****.*

26.27.
Cette définition demande quelques remarques.

(1) La notion de base hilbertienne 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 ℓ^2.

(3) Nous ne demandons pas que la famille soit libre ? La belle affaire. Une famille orthogonale est toujours libre, proposition 26.23.

Lemme 26.28 ([388])**.**
Si $\mathcal{E}$ est une base hilbertienne[9] *de l'espace de Hilbert $\mathscr{H}$ et si $x \in \mathscr{H}$ alors l'ensemble*

$$\{e \in \mathcal{E} \text{ tel que } \langle x, e\rangle \neq 0\} \tag{26.68}$$

est au plus dénombrable.

Démonstration. Notons qu'ici, $\mathscr{H}$ n'est pas supposé séparable. Nous savons par l'inégalité de Bessel (26.61) que

$$\sum_{e\in\mathcal{E}} |\langle x, e\rangle|^2 \leqslant \|x\|^2. \tag{26.69}$$

Donc si $\epsilon > 0$ est donné, l'ensemble $\{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > \epsilon\}$ est fini. Or

$$\{e \in \mathcal{E} \text{ tel que } \langle x, e\rangle \neq 0\} = \{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > 0\} = \bigcup_{n=1}^{\infty} \{e \in \mathcal{E} \text{ tel que } |\langle x, e\rangle| > \frac{1}{n}\}. \tag{26.70}$$

Bref, cet ensemble est une union dénombrable d'ensembles finis. Il est donc dénombrable. □

Remarque 26.29.
Le lemme 26.28 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.

Corollaire 26.30 ([388])**.**
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 $\mathscr{H}$, 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 (26.66) de f dans la base $\mathcal{E}$, nous aurions $f = 0$. Nous avons donc

$$\mathcal{F} \subset \bigcup_{e\in\mathcal{E}} \{f \in \mathcal{F} \text{ tel que } \langle f, e\rangle \neq 0\}, \tag{26.71}$$

alors que le lemme 26.28 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. □

8. Définition 26.19
9. Définition 26.26.

Remarque 26.31.
En travaillant un peu plus sur la notion de cardinalité, le corollaire 26.30 indique que toutes les bases hilbertiennes ont même cardinalité. En effet en laissant tomber l'hypothèse de dénombrabilité sur $\mathcal{E}$, l'inclusion (26.71) donne que $\mathcal{F}$ est une union de $\text{Card}(\mathcal{E})$ ensembles dénombrables et est alors de cardinalité $\text{Card}(\mathcal{E})$.

Définition 26.32.
*Une partie orthonormale B est **maximale** si le seul $x \in \mathscr{H}$ vérifiant $\langle x, b\rangle = 0$ pour tout $b \in B$ est $x = 0$.*

Lemme 26.33 ([1, 389])**.**
Tout espace de Hilbert possède une partie orthonormale maximale.

Démonstration. Soit $\mathscr{H}$ un espace de Hilbert et $\mathcal{O}$, l'ensemble de parties orthonormales de $\mathscr{H}$, ordonné [10] 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 $\mathscr{H}$, 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

$$A = \bigcup_{i\in I} A_i. \tag{26.72}$$

Cela est encore une partie orthonormé de $\mathscr{H}$ 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 1.15 nous dit alors que $\mathcal{O}$ possède un maximum. Ce maximum est une partie orthonormale inclue dans aucune autre partie orthonormale. C'est-à-dire que ce maximum est une partie orthonormale maximale au sens de la définition 26.32. □

Proposition 26.34 ([389])**.**
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

$$A(b) = \{a \in A \text{ tel que } \langle a, b\rangle \neq 0\} \cup \{0\}. \tag{26.73}$$

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

$$A \subset \bigcup_{b\in B} A(b). \tag{26.74}$$

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. □

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 26.36, vous pouvez sauter la 26.35.

Proposition 26.35.
Tout espace de Hilbert séparable possède une base hilbertienne.

10. Définition 1.7.

Démonstration. Vu que nous supposons avoir un espace de Hilbert séparable, il possède une partie totale dénombrable par la proposition 26.20. 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

$$F_n = \text{Span}\{v_1, \dots, v_n\}. \tag{26.75}$$

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 Span ne considère que des combinaisons linéaires finies. □

Proposition 26.36 ([389])**.**
À propos de parties orthonormales maximales.

(1) *Une partie d'un espace de Hilbert est orthonormale maximale si et seulement si elle est une base hilbertienne.*

(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 $F = \mathscr{H}$. Pour cela nous considérons $x \in \mathscr{H}$ et nous utilisons le théorème de projection orthogonale 26.7 pour mentionner le fait que

$$x - \texttt{proj}_F(x) \perp F \tag{26.76}$$

Cela dit que $x - \texttt{proj}_F(x)$ est un vecteur orthogonal en particulier à tous les éléments de B ; par maximalité nous avons $x - \texttt{proj}_F(x) = 0$, c'est-à-dire $x = \texttt{proj}_F(x)$ ou encore $x \in F$. Cela prouve que $F = \mathscr{H}$.

Note : pour être pointilleux, nous aurions dû travailler non avec $x - \texttt{proj}_F(x)$, mais avec le vecteur normalisé à 1.

En ce qui concerne le second point, nous invoquons le lemme 26.33 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. □

Voici un petit résumé de ce que nous avons vu en termes de dénombrabilité et séparabilité.

Théorème 26.37.
Pour un espace de Hilbert, les choses suivantes sont équivalentes.

(1) *L'espace est séparable*[11].

(2) *L'espace possède au moins une base hilbertienne*[12] *dénombrable.*

(3) *Toutes les bases hilbertiennes sont dénombrables.*

(4) *Toute partie libre est dénombrable.*

Démonstration. Plein de résultats à citer ...

(1) implique (2) est la proposition 26.35.

(2) implique (3) est la proposition 26.30.

(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 26.34.

11. Définition 26.18.
12. Définition 26.26.

(4) implique (1) La proposition 26.36(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 26.23), 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 26.20.

□

Remarque 26.38.
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 26.24 en disant que la somme sur la « partie de base » est plus petite que la somme sur la « base complète ».

26.39.
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 indicer 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 26.40.
Soit $(u_k)_{k\geqslant 1}$ un système orthonormé d'un préhilbert $\mathscr{H}$. Soient

$$x = \sum_{k=1}^{\infty} \xi_k u_k \tag{26.77}$$

et

$$F = \text{Span}\{u_1, \dots, u_n\}. \tag{26.78}$$

Alors

$$\texttt{proj}_F(x) = \sum_{k=1}^{n} \xi_k u_k. \tag{26.79}$$

Démonstration. Nous allons dans un premier temps montrer que

$$y = x - \sum_{k=1}^{n} \xi_k u_k \tag{26.80}$$

est dans $F^\perp$. Pour cela nous calculons

$$\langle x - \sum_{k=1}^{n} \xi_k u_k, u_j, \rangle = \langle \sum_{k=n+1}^{\infty} \xi_k u_k, u_j \rangle = 0 \tag{26.81}$$

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 $\texttt{proj}_F y = 0$ par le point (2) du théorème 26.12.

D'autre part $\texttt{proj}_F y$ peut être calculé selon

$$\texttt{proj}_F y = \texttt{proj}_F x - \sum_{k=1}^{n} \xi_k \texttt{proj}_F u_k \tag{26.82}$$

tandis que $\texttt{proj}_F u_k = u_k$ lorsque $1 \leqslant k \leqslant n$. Par conséquent l'annulation de $\texttt{proj}_F y$ donne

$$\texttt{proj}_F x = \sum_{k=1}^{n} \xi_k u_k, \tag{26.83}$$

donc le résultat. □

Théorème 26.41 (Meilleur approximation)**.**
Soit $\{u_i\}_{i\in I}$ une famille orthonormé de l'espace de Hilbert $\mathscr{H}$. Alors pour tout $x \in \mathscr{H}$ et pour tout J fini dans I et pour toute famille de nombres complexes $(a_j)_{j\in J}$ nous avons

$$\| \sum_{j\in J} \langle x, u_j\rangle u_j - x\| \leqslant \| \sum_{j\in J} a_j u_j - x\|. \tag{26.84}$$

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.

Corollaire 26.42.
Soit $\{u_i\}_{i\in I}$ une famille orthonormé de $\mathscr{H}$. Pour tout $x \in \mathscr{H}$ et pour toutes parties finies J, K de I avec $J \subset K$ nous avons

$$\| \sum_{j\in K} \langle x, u_j\rangle u_j - x\| \leqslant \| \sum_{j\in J} \langle x, u_j\rangle u_j - x\|. \tag{26.85}$$

Ce corollaire exprime le fait que plus on prend de termes de la forme $\langle , u_i\rangle u_i$, mieux c'est.

Proposition 26.43.
Soit $\mathscr{H}$ un espace de Hilbert et (u_n) un système orthonormé dans $\mathscr{H}$. Si $(\xi_n)_{n\in\mathbb{N}}$ est une suite dans ℓ^2 alors la série

$$\sum_{n=1}^{\infty} \xi_n u_n \tag{26.86}$$

converge dans $\mathscr{H}$.

Autrement dit l'application

$$\begin{aligned} S\colon \mathscr{H} &\to \ell^2 \\ x &\mapsto \big(\langle x, u_n\rangle\big)_{n\geqslant 1} \end{aligned} \tag{26.87}$$

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

$$\lim_{n\to\infty} \| \sum_{k=n}^{n+p} \xi_k u_k\| = 0 \tag{26.88}$$

est uniforme en p. Cela est un corollaire de la formule (26.53) parce que

$$\| \sum_{k=n}^{n+p} \xi_k u_k\|^2 = \sum_{k=n}^{n+p} |x_k|^2. \tag{26.89}$$

Mais si (ξ_n) est dans ℓ^2, pour tout ϵ, il existe N tel que si $n > N$ alors le membre de droite est inférieur à ϵ indépendamment de p. □

26.3.5 Décomposition dans une base hilbertienne

Étant donné une base hilbertienne de $\mathscr{H}$, nous notons

$$c_k(x) = \langle x, u_k\rangle. \tag{26.90}$$

Dans le théorème suivant (et d'ailleurs partout), les sommes sur I sont prises au sens de la définition 12.100.

Théorème 26.44 (Décomposition dans une base orthogonale)**.**
Soit $\mathscr{H}$ un espace de Hilbert séparable $\{u_i\}_{i\in I}$ une base orthonormée (I est un ensemble dénombrable quelconque).

(1) *Pour tout $x \in \mathscr{H}$ nous avons*

$$x = \sum_{i\in I} \langle x, u_i\rangle u_i \tag{26.91}$$

où la somme est prise au sens de la définition 12.100. En particulier, la somme converge de façon commutative.

(2) *Si $\{e_i\}_{i\in I}$ est une famille orthonormée qui satisfait la décomposition (26.91) pour tout $x \in \mathscr{H}$ alors $\{e_i\}$ est une base hilbertienne.*

(3) *Nous avons l'identité de* ***Plancherel***

$$\|x\|^2 = \sum_{i\in I} |\langle x, u_i\rangle|^2. \tag{26.92}$$

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***

$$\langle x, y\rangle = \sum_{i\in I} \langle x, u_i\rangle\overline{\langle y, u_i\rangle}. \tag{26.93}$$

(5) *Si $\{e_i\}$ est une famille de vecteurs unitaires vérifiant l'identité de Plancherel, alors c'est une base hilbertienne.*

(6) *Si $\{u_i\}_{i\in I}$ est une base hilbertienne, la suite $n \mapsto |\langle x, e_n\rangle|$ 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

$$x_n = \sum_{x\in J_n} a_{n,k} u_j \tag{26.94}$$

et $x_n \to x$ dans $\mathscr{H}$. 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 $n \geqslant N$. 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

$$\sum_{j\in J_N} \langle x, u_j\rangle u_j = \sum_{j\in J} a_j u_j \tag{26.95}$$

avec

$$a_j = \begin{cases} \langle x, u_j\rangle & \text{si } j \in J_N \\ 0 & \text{sinon,} \end{cases} \tag{26.96}$$

le théorème de meilleure approximation 26.41 nous enseigne que

$$\| \sum_{j\in J} \langle x, u_j\rangle u_j - x\| \leqslant \| \sum_{j\in J_N} \langle x, u_j\rangle u_j - x\| < \epsilon \tag{26.97}$$

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 $\mathscr{H}$. Par hypothèse, pour tout ϵ, il existe un ensemble fini J tel que

$$\| \sum_{j\in J} \langle x, u_j\rangle u_j - x\| < \epsilon. \tag{26.98}$$

Cela prouve la densité dont nous avions besoin.

(3) La norme étant une fonction continue, elle commute avec les sommes infinies, de telle sorte que l'égalité de Plancherel donne

$$\|x\|^2 = \|\sum_{i\in I}\langle x,u_i\rangle u_i\|^2. \tag{26.99}$$

Le système des u_i étant orthonormé,

$$\|x\|^2 = \|\sum_{i\in I}\langle x,u_i\rangle u_i\|^2 = \sum_{i\in I}|\langle x,u_i\rangle|^2. \tag{26.100}$$

(4) Au tour de Parseval. Nous commençons par prouver que la somme du membre de droite converge. En utilisant l'inégalité $|zz'| \leqslant |z|^2 + |z'|^2$ (valable pour $z,z' \in \mathbb{C}$) et Plancherel, nous avons

$$\begin{aligned}\sum_{i\in I}|\langle x,u_i\rangle\overline{\langle y,u_i\rangle}| &\leqslant \sum_{i\in I}|\langle x,u_i\rangle|^2 + |\langle y,u_i\rangle|^2\\ &\leqslant \|x\|^2 + \|y\|^2.\end{aligned} \tag{26.101}$$

Nous en déduisons que la famille $\langle x,u_i\rangle\overline{\langle y,u_i\rangle}$ est (commutativement) sommable en utilisant la proposition 12.107. Par ailleurs nous savons que

$$x = \sum_{i\in I}c_i(x)u_i \tag{26.102a}$$

$$y = \sum_{i\in I}c_i(y)u_i, \tag{26.102b}$$

et le produit scalaire étant une forme bilinéaire continue,

$$\langle x,y\rangle = \sum_{i\in I}\sum_{j\in I}\langle c_i(x)u_i, \overline{c_j(y)u_j}\rangle \tag{26.103a}$$

$$= \sum_{i\in I}\sum_{j\in J}c_i(x)\overline{c_j(y)}\delta_{ij} \tag{26.103b}$$

$$= \sum_{i\in I}c_i(x)\overline{c_j(y)}. \tag{26.103c}$$

(5) Nous utilisons l'égalité de Plancherel avec $x = u_j$:

$$\|u_j\|^2 = \|u_j\| + \sum_{i\in I\setminus\{j\}}|\langle u_j,u_i\rangle|^2. \tag{26.104}$$

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 (26.64) que nous avions déduites dans la démonstration de l'inégalité de Bessel :

$$\|x - \sum_{j\in J}c_j(x)u_j\| = \|x\|^2 - \sum_{j\in J}|c_j(x)|^2. \tag{26.105}$$

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 $x \in \mathscr{H}$ 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 26.45.
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 26.46.
L'égalité de Parseval est la raison pour laquelle les physiciens écrivent souvent

$$\mathrm{Id} = \sum_{n=1}^{\infty} |u_n\rangle\langle u_n| \tag{26.106}$$

dans les livres de mécanique quantique par exemple. Dans certains[390], nous lisons même

$$\int_{-\infty}^{+\infty} dq|q\rangle\langle q| = \hat{I}. \tag{26.107}$$

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 26.47.
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 26.48
L'identité de Parseval (26.93), dans le cas de l'espace des fonctions continues périodiques de période 2π signifie qu'en posant

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(s)e^{-ins}, \tag{26.108}$$

nous avons

$$\frac{1}{2\pi}\int_0^{2\pi} |f(s)|^2 = \sum_{n=-\infty}^{\infty} |c_n(f)|^2. \tag{26.109}$$

△

Corollaire 26.49.
Soit $\mathscr{H}$ un espace de Hilbert et (u_n) une base orthonormée. L'application

$$\begin{aligned} S\colon \mathscr{H} &\to \ell^2 \\ x &\mapsto \big(\langle x, u_n\rangle\big)_{n\in\mathbb{N}} \end{aligned} \tag{26.110}$$

est un isomorphisme d'espaces de Hilbert.

De plus l'isomorphisme réciproque est

$$\begin{aligned} S^{-1}\colon \ell^2 &\to \mathscr{H} \\ (\xi_n) &\mapsto \sum_{n=1}^{\infty} \xi_n u_n. \end{aligned} \tag{26.111}$$

Démonstration. Nous devons prouver que l'application est bijective et qu'elle vérifie

$$S(x) \cdot S(y) = \langle x, y\rangle \tag{26.112}$$

où le point dénote le produit dans ℓ^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 26.43 et par conséquent (ξ_n) est l'image par S de ce vecteur de $\mathscr{H}$.

Pour l'injectivité, si $S(x) = S(y)$ alors

$$x = \sum_n \langle x, u_n\rangle u_n = \sum_n \langle y, u_n\rangle u_n = y \tag{26.113}$$

en utilisant la décomposition (26.91).

Le fait que S soit une isométrie est contenu dans Parseval. □

Proposition 26.50 ([1]).
Soit un espace de Hilbert séparable $\mathscr{H}$, un sous-espace vectoriel fermé V et une base orthonormée $\{b_i\}_{i\in I}$ de $\mathscr{H}$. En posant $v_i = \texttt{proj}_V(b_i)$ alors

$$C = \overline{\text{Span}\{b_i - v_i\}} \tag{26.114}$$

est l'orthogonal de V.

Démonstration. Juste pour rappel, lorsque nous écrivons $v_i = \texttt{proj}_V(b_i)$, nous parlons de la projection orthogonale du théorème 26.7. 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 $c_i = b_i - v_i = b_i - \texttt{proj}_V(b_i)$. Alors d'une part

$$0 = \texttt{proj}_V(x) = \sum_i x_i \texttt{proj}_V(b_i) = \sum_i x_i(b_i - c_i). \tag{26.115}$$

Nous avons utilisé la continuité de $\texttt{proj}_V$ pour permuter avec la somme. D'autre part,

$$x = \sum_i x_i b_i = \sum_i x_i(b_i - c_i + c_i) = \underbrace{\sum_i x_i(b_i - c_i)}_{=0} + \sum_i x_i c_i \in C. \tag{26.116}$$

Notons que pour la dernière appartenance, il est important de prendre la fermeture pour définir C. □

Proposition 26.51 ([1, 391]).
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 $\mathscr{H}$. Nous mentionnons que cette partie est libre et donc, par le théorème 26.37, dénombrable [13]. Montrons pour commencer que la partie $V = \overline{\text{Span}\{u_i\}_{i\in I}}$ est un sous-espace vectoriel fermé de $\mathscr{H}$.

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,

$$\|v + w - (a + b)\| \leqslant 2\epsilon. \tag{26.117}$$

Cela prouve que $v + w \in V$. Nous procédons de même pour λv.

Fermé Par construction.

L'orthogonal [14] de V est un sous-espace vectoriel de $\mathscr{H}$, 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 [15] de $\mathscr{H}$.

$C \cup U$ est libre Bing ! Il ne faut pas le démontrer : ça ne fait pas partie de la définition d'une base hilbertienne.

$C \cup U$ est générateur Bang ! Il ne faut pas le démontrer : ça ne fait pas partie de la définition d'une base hilbertienne.

$C \cup U$ 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 $V^\perp$.

13. 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.
14. Voir la définition 26.9 et la proposition 26.50.
15. Définition 26.26.

$\underline{C \cup U \text{ est dénombrable}}$ 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.

$\underline{C \cup U \text{ est total}}$ Nous devons prouver que $\overline{\text{Span}(C \cup U)} = \mathscr{H}$ parce qu'il y a bien la fermeture qui intervient dans la définition 26.19. Pour cela nous utilisons la proposition 26.14. Si $x \in \mathscr{H}$ alors nous avons

$$x = \texttt{proj}_V(x) + \big(x - \texttt{proj}_V(x)\big) = \sum_{i\in I} x_i u_i + \sum_{\alpha\in A} x_\alpha c_\alpha \tag{26.118}$$

pour des coefficients x_i et x_α. Notons que I et A sont deux ensembles différents. Aucun des x_α n'est un des x_i, ni inversement. Supposons que $x \in \text{Span}(C \cup U)^\perp$; alors

$$0 = \langle x, u_j\rangle = \sum_{i\in I} x_i \underbrace{\langle u_i, u_j\rangle}_{=\delta_{ij}} + \sum_{\alpha\in A} \underbrace{\langle c_\alpha, u_j\rangle}_{=0} = x_j. \tag{26.119}$$

Donc $x_j = 0$. En faisant de même avec $0 = \langle x, c_\beta\rangle = x_\beta$ nous déduisons $x = 0$.

□

Proposition 26.52 ([1])**.**
Si $\mathscr{H}$ est un espace de Hilbert séparable de base hilbertienne $\{e_i\}_{i\in I}$, alors $\mathscr{H}'$ est un espace de Hilbert séparable dont une base hilbertienne est donnée par les formes

$$\begin{aligned} \alpha_i \colon \mathscr{H} &\to \mathbb{C} \\ e_j &\mapsto \delta_{ij}. \end{aligned} \tag{26.120}$$

Démonstration. C'est la proposition 26.4 qui fait tout. □

26.3.6 Digression sur les normes opérateurs

Le théorème 26.44 nous indique que si $\{u_i\}_{i\in I}$ est une base hilbertienne, alors pour tout $x \in \mathscr{H}$ nous avons

$$\texttt{proj}_{u_i}(x) = \langle x, u_i\rangle u_i, \tag{26.121}$$

et donc

$$\sum_{i\in I} \texttt{proj}_{u_i} x = x. \tag{26.122}$$

Nous ne pouvons cependant pas conclure que

$$\sum_{i\in I} \texttt{proj}_{u_i} = \text{Id} \tag{26.123}$$

au sens de la norme opérateur de la définition 12.10. En effet en prenant $I = \mathbb{N}$, l'égalité (26.123) demanderait d'avoir

$$\lim_{N\to\infty} \left\| \sum_{i=1}^{N} \texttt{proj}_{u_i} - \text{Id} \right\|_\infty = 0, \tag{26.124}$$

or pour tout N, le vecteur u_{N+1} réalise

$$\sum_{i=1}^{N} \texttt{proj}_{u_i}(u_{N+1}) - u_{N+1} = -u_{N+1}. \tag{26.125}$$

Par conséquent pour tout N nous avons

$$\sup_{\|x\|=1} \left\| \sum_{i=1}^{N} \texttt{proj}_{u_i} x - x \right\| \geqslant 1. \tag{26.126}$$

Nous ne pouvons donc pas dire

$$\sum_{n=1}^{\infty} \mathtt{proj}_{u_i} = \mathrm{Id} \tag{26.127}$$

au sens de la norme opérateur.

Nous avons cependant la convergence au sens faible.

Proposition 26.53.
Soit $\mathscr{H}$ un espace de Hilbert et $\{u_i\}_{i\in\mathbb{N}}$ une base hilbertienne de $\mathscr{H}$. Au sens de la topologie faible sur l'espace des opérateurs nous avons

$$\sum_{i=1}^{\infty} \mathit{proj}_{u_i} = \mathrm{Id}\,. \tag{26.128}$$

Démonstration. Pour chaque $N \in \mathbb{N}$ et $x \in \mathscr{H}$, en vertu de la décomposition (26.91) nous avons

$$\sum_{i=1}^{N} \mathtt{proj}_{u_i}(x) - x = \sim_{i=N+1}^{\infty} \langle x, u_i\rangle u_i. \tag{26.129}$$

Par l'orthonormalité de la base nous avons

$$\sum_{i=N+1}^{\infty} \|\langle x, u_i\rangle u_i\| = \sum_{i=N+1}^{\infty} |\langle x, u_i\rangle|, \tag{26.130}$$

dont la limite $N \to \infty$ est zéro étant donné que la suite $i \mapsto |\langle x, u_i\rangle|$ est dans $\ell^2(\mathbb{R})$ par le théorème 26.44. □

26.3.7 Applications linéaires et continuité

Nous avons déjà vu dans l'exemple 12.34 que la fonction

$$\begin{aligned} f\colon H &\to H \\ e_k &\mapsto k e_k \end{aligned} \tag{26.131}$$

n'était pas continue en zéro alors qu'elle est linéaire. Nous allons maintenant voir qu'elle est un contre-exemple à la proposition 13.223. Calculons les dérivées partielles :

$$\frac{\partial f}{\partial e_j}(x) = \frac{d}{dt}\Big[f(x + te_j)\Big]_{t=0} = \frac{d}{dt}\Big[\sum_k k(x_k + t\delta_{jk})e_k\Big]_{t=0} = je_j. \tag{26.132}$$

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 13.223 est l'introduction des « contres-termes » dans l'équation (13.570). En effet les contre-termes à ajouter seraient

$$l(x) = \frac{d}{dt}\Big[f\big(a + s(x - a)\big)\Big]_{t=0} \tag{26.133}$$

qui ici serait

$$l\big(\sum_k x_k e_k\big) = \frac{d}{dt}\Big[f\big(s\sum_k x_k e_k\big)\Big]_{t=0} = \sum_k x_k e_k, \tag{26.134}$$

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

$$\frac{\partial f}{\partial u}(x) = \frac{d}{dt}\Big[f(x + tu)\Big]_{t=0} = \frac{d}{dt}\Big[\sum_k k(x_k + tu_k)e_k\Big]_{t=0} = \sum_k k u_k e_k \tag{26.135}$$

La convergence de la dernière somme n'est pas garantie pour tout u.

Exemple 26.54
Soit un espace de Hilbert $\mathscr{H}$. L'application

$$\begin{aligned}\phi\colon \mathscr{H} &\to \mathbb{C}\\ e_i &\mapsto 1\end{aligned} \tag{26.136}$$

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

$$\|y_n\|^2 = \langle y_n, y_n\rangle = \sum_{i,j}^{n}\langle e_i, e_j\rangle = n, \tag{26.137}$$

donc $\|y_n\| = \sqrt{n}$, et

$$\phi(x_n) = \frac{n}{\sqrt{n}} = \sqrt{n}. \tag{26.138}$$

L'application ϕ n'est donc pas bornée et pas continue non plus. △

Proposition 26.55.
Soient deux espaces de Hilbert séparables[16] $\mathscr{H}_1$ *et* $\mathscr{H}_2$. *Soit une base orthonormée* $\{e_i\}_{i\in I}$ *où* I *est dénombrable. Si* $\Phi\colon \mathscr{H}_1 \to \mathscr{H}_2$ *est une bijection linéaire continue, alors*

$$\Phi\big(\sum_{i\in I} x_i e_i\big) = \sum_{i\in I} x_i \Phi(e_i). \tag{26.139}$$

Démonstration. Soit $\epsilon > 0$. Vu que $\sum_{i\in I} x_i e_i$ est une somme convergente[17], il existe une partie finie J de I telle que pour tout K fini contenant J,

$$\|\sum_{j\in K} x_j e_j - x\| < \epsilon. \tag{26.140}$$

Vu que Φ est une isométrie linéaire, nous avons

$$\|\sum_{j\in K} x_j \Phi(e_j) - \Phi(x)\| < \epsilon, \tag{26.141}$$

et donc bien l'égalité

$$\sum_{i\in I} x_i \Phi(e_j) = \Phi(x). \tag{26.142}$$

□

La proposition suivante donne une formule pour l'inverse du Φ du théorème de représentation de Riesz 26.17 lorsque nous avons une base hilbertienne.

Proposition 26.56.
Soit un espace de Hilbert séparable $\mathscr{H}$. *Si* $\phi \in \mathscr{H}'$, *alors*

$$\Phi^{-1}(\phi) = \sum_{i\in I} \phi(e_i) e_i \tag{26.143}$$

dès que $\{e_i\}_{i\in I}$ *est une base hilbertienne de* $\mathscr{H}$.

16. Donc à base dénombrable, voir 26.37.
17. Définition 12.100.

Démonstration. Nous considérons la base des $\{\alpha_i\}_{i\in I}$ de $\mathscr{H}'$. 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 26.55 pour permuter Φ et la somme nous avons le calcul suivant :

$$\phi = \sum_{i\in I}\phi_i\alpha_i = \sum_{i\in I}\phi_i\Phi(e_i) = \sum_{i\in I}\Phi(\phi_i e_i) = \Phi\big(\sum_{i\in I}\phi_i e_i\big) = \Phi(\sum_{i\in I}\phi(e_i)e_i). \tag{26.144}$$

En résulté :

$$\phi = \Phi(\sum_{i\in I}\phi(e_i)e_i), \tag{26.145}$$

et donc la formule (26.143). □

26.4 Théorème de Kochen-Specker

Le théorème suivant est central en mécanique quantique. La démonstration provient de [392] et de Wikipédia. Nous allons démontrer complètement le théorème seulement pour les espaces de Hilbert de dimension plus grande ou égale à 4.

Théorème 26.57 (Kochen-Specker[392])**.**
Soit $\mathscr{H}$ un espace de Hilbert de dimension plus grande ou égale à 3. Une fonction v sur l'ensemble des opérateurs de $\mathscr{H}$ 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 $\mathscr{H}$. Nous notons $\texttt{proj}_i$ 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 [18].

D'abord nous devons avoir $v(\mathbb{1}) = 1$. En effet pour tout opérateur A, nous avons

$$v(A) = v(A\mathbb{1}) = v(A)v(\mathbb{1}). \tag{26.146}$$

Pour peu que $v(A) \neq 0$, cela nous fait $v(\mathbb{1}) = 1$.

En vertu du théorème 26.44, un vecteur $x \in \mathscr{H}$ se décompose en $x = \sum_n\langle x, u_n\rangle u_n$, et nous avons

$$\texttt{proj}_i\, x = \langle x, u_i\rangle u_i. \tag{26.147}$$

En effet le théorème de la projection orthogonale 26.7 nous enseigne que $\texttt{proj}_i\, x$ serait l'unique vecteur de la forme λu_i tel que $\lambda u_i - x \perp u_i$. Il est facile de vérifier que le vecteur proposé par (26.147) vérifie cette propriété.

Une conséquence est que

$$\left(\sum_{i=0}^{\infty}\texttt{proj}_i\right)(x) = x. \tag{26.148}$$

Par conséquent, par hypothèse du théorème nous devons avoir

$$\sum_i v(\texttt{proj}_i) = v(\mathbb{1}) = 1. \tag{26.149}$$

Étant donné que les projections sont idempotentes,

$$v(\texttt{proj}_i) = v(\texttt{proj}_i^2) = v(\texttt{proj}_i)^2 \tag{26.150}$$

18. Pour les besoins de la physique, nous remarquons que ces opérateurs sont des opérateurs hermitiens qui commutent, mais ça ne joue pas ici.

et donc $v(\texttt{proj}_i)$ doit valoir zéro ou un. Mais la relation (26.149) donne une forte contrainte sur le choix de 0 et de 1. En effet, parmi les $v(\texttt{proj}_i)$, 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 $v(\texttt{proj}_i)$ revient à attacher 0 ou 1 à tous les *rayons* de $\mathscr{H}$ de façon que toute base orthogonale de $\mathscr{H}$ 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 $\mathscr{H}$ 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

$$\{Ru_0, Ru_1, Ru_2, Ru_3, u_k\}_{k\geqslant 3} \tag{26.151}$$

est encore une base orthogonale de $\mathscr{H}$ et nous avons encore $v(u_k) = 0$ pour $k \geqslant 4$. 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 base 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 écris 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.

□

26.5 Théorème de Lax-Milgram

Définition 26.58.
Une forme bilinéaire $a\colon V \times V \to \mathbb{R}$ sur un espace vectoriel normé V est ***coercitive*** *s'il existe $\alpha > 0$ tel que $a(u, u) \geqslant \alpha\|u\|^2$ pour tout $u \in V$.*

Théorème 26.59 (Lax-Milgram[393])**.**
Soit un espace de Hilbert réel V muni de différentes choses.

(1) L'application linéaire $L\colon V \to \mathbb{R}$ 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[19]*.*

19. Définition 26.58.

Alors

(1) *Il existe $M > 0$ tel que $|a(u,v)| \leqslant M\|u\|\|v\|$ 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*

$$\|u\| \leqslant \frac{\|L\|}{\alpha} \tag{26.152}$$

Démonstration. Évacuons deux points faciles avant de commencer les choses sérieuses. D'abord L est continue par la proposition 12.25. Ensuite, l'existence du M est donnée par la proposition 12.43.

Maintenant nous commençons.

Reformulation en équation linéaire La forme L est continue et donc dans le dual V' ; le théorème de Riesz 26.17 nous donne donc $f \in V$ tel que

$$L(v) = \langle f, v \rangle \tag{26.153}$$

pour tout $v \in V$. De plus si $w \in V$ est fixé, l'application $b_w \colon v \mapsto a(w,v)$ est linéaire et bornée parce que

$$\|b_w\| = \sup_{\|v\|=1} |b_w(v)| = \sup_{\|v\|=1} |a(w,v)| \leqslant M\|w\|\|v\| = M\|w\|. \tag{26.154}$$

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

$$b_w(v) = \langle A(w), v \rangle \tag{26.155}$$

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

$$A(u) = f. \tag{26.156}$$

A est linéaire Soient $\alpha, \beta \in \mathbb{R}$ et $w, z \in V$. Nous avons pour tout $v \in V$:

$$\begin{aligned}
\langle A(\alpha w + \beta z), v\rangle &= a(\alpha w + \beta z, v) && \text{(26.157a)}\\
&= \alpha a(w,v) + \beta a(z,v) && \text{(26.157b)}\\
&= \alpha\langle A(w), v\rangle + \beta\langle A(z), v\rangle && \text{(26.157c)}\\
&= \langle \alpha A(w) + \beta A(z), v\rangle. && \text{(26.157d)}
\end{aligned}$$

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.

Une autre propriété de A Nous déduisons une majoration de $\|A(v)\|^2$ lorsque ce n'est pas nul. Pour ce faire,

$$\begin{aligned}
\|A(v)\|^2 a &= \langle A(v), A(v)\rangle && \text{(26.158a)}\\
&= a(v, A(v)) && \text{(26.158b)}\\
&\leqslant M\|v\|\|A(v)\|. && \text{(26.158c)}
\end{aligned}$$

En simplifiant, $\|A(v)\| \leqslant M\|v\|$. Et donc

$$\|A(v)\|^2 \leqslant M^2\|v\|^2. \tag{26.159}$$

Une contraction Nous allons choisir une valeur de $\rho > 0$ telle que l'application

$$T \colon w \mapsto w - \rho\big(A(w) - f\big) \tag{26.160}$$

soit une contraction [20]. Nous avons $T(w)-T(w')=w-w'-\rho\big(A(w-w')\big)$ et donc

$$\|T(w)-T(w')\|^2=\|w-w'\|^2+\rho^2\|A(w-w')\|^2-2\rho\langle A(w-w'),w-w'\rangle \tag{26.161a}$$
$$=\|w-w'\|^2+\rho^2\|A(w-w')\|^2-2\rho a(w-w',w-w'). \tag{26.161b}$$

Vu que le dernier terme arrive avec un signe moins, pour majorer l'expression, il faut minorer ce terme, c'est-à-dire utiliser $a(w-w',w-w")\geqslant\alpha\|w-w-w'\|$. Et en même temps nous utilisons (26.159) pour le second terme. Au final pour pouvons factoriser $\|w-w'\|$ et

$$\|T(w)-T(w')\|\leqslant\|w-w'\|\big(1+\rho^2M^2-2\rho\alpha\big). \tag{26.162}$$

Pout que T soit contractante, il faut $0<P(x)<1$ avec $P(x)=M^2x^2-2\alpha x+1$. 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 $\lim_{x\to\infty}P(x)=+\infty$, 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 ρ cette valeur, l'application T est une contraction.

Point fixe et conclusion L'ensemble des solutions du problème (26.156) est égal à l'ensemble des points fixes de $T(v)=v-\rho\big(A(v)-f\big)$.

L'application $T\colon V\to V$ est contractante et V est métrique et complet. Ergo le théorème de point fixe de Picard 18.34 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.

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

$$a(u,u)=L(u). \tag{26.163}$$

D'un côté nous utilisons $a(u,u)\geqslant\alpha\|u\|^2$ et de l'autre, $L(u)\leqslant\|L\|\|u\|$:

$$\alpha\|u\|^2\leqslant\|L\|\|u\| \tag{26.164}$$

et donc

$$\|u\|\leqslant\frac{\|L\|}{\alpha}. \tag{26.165}$$

Notons que $L(u)$ et $a(u,u)$ sont positifs.

□

Exemple 26.60([394])
La borne $\|u\|\leqslant\|L\|/\alpha$ est optimale au sens où il existe des cas d'égalité. En effet nous pouvons considérer

$$a(u,v)=s\langle u,v\rangle \tag{26.166}$$

et

$$L(v)=\langle f,v\rangle \tag{26.167}$$

pour un certain $s>0$ et $f\in V$. L'application L n'est autre que Φ_f dont nous avons abondamment parlé autour du théorème de représentation de Riesz 26.17. Le lemme 26.15 nous apprend que L est une application borné (et donc continue) de norme $\|L\|_{V'}=\|f\|_V$.

L'application $a\colon V\times V\to\mathbb{R}$ est continue par la proposition 12.47. Elle est coercive parce que

$$a(u,u)=s\|u\|^2. \tag{26.168}$$

Les hypothèses du théorème de Lax-Milgram 26.59 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

$$u=\frac{1}{s}f. \tag{26.169}$$

20. Définition 18.33.

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 Φ est une isométrique dans le lemme 26.15. En reprenant les notations du théorème, nous avons l'égalité

$$\|u\| = \frac{\|L\|}{\alpha}, \tag{26.170}$$

ce qui signifie que l'inégalité du théorème est dans un certain sens optimale. △

Proposition 26.61.
Soit un espace de Hilbert réel V et une application $a\colon V \times V \to \mathbb{R}$ coercive et continue.

L'application $\phi\colon V' \to V$ 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 ϕ soit bien définie est le théorème de Lax-Milgram 26.59.

L'application ϕ 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

$$a(u_1, v) = L_1(v) \tag{26.171a}$$
$$a(u_2, v) = L_2(v) \tag{26.171b}$$

pour tout $v \in V$. Par bilinéarité de a nous avons aussi

$$a(u_1 + u_2, v) = (L_1 + L_2)(v). \tag{26.172}$$

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 $\lambda \in \mathbb{R}$,

$$a(\lambda u, v) = (\lambda L)(v), \tag{26.173}$$

c'est-à-dire que $\phi(\lambda L) = \lambda\phi(L)$.

Donc ϕ est linéaire. Pour la continuité, il suffit de s'assurer que ϕ est bornée. Pour cela nous utilisons la majoration (26.152) et la définition de la norme d'une application linéaire :

$$\|\phi\| = \sup_{\|L\|=1} \|\phi(L)\| \leqslant \sup_{\|L\|=1} \frac{1}{\alpha}\|L\| = \frac{1}{\alpha}. \tag{26.174}$$

Donc ϕ est bornée et par voie de conséquence continue (proposition 12.25). □

Un problème est bien posé au sens de Hadamard si la solution est unique et dépend continuement des données ; nous ne rentrerons pas plus dans les détails pour l'instant : nous en reparlerons dans la définition 34.14 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 26.61 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 26.62 (Lax-Milgram version symétrique[393])**.**
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 $L\colon V \to \mathbb{R}$ 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 $|a(u,v)| \leqslant M\|u\|\|v\|$ 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 elle minimise la ***fonctionnelle d'énergie***

$$J(w) = \frac{1}{2}a(w,w) - l(w). \tag{26.175}$$

Démonstration. Nous séparons les deux sens.

$\Rightarrow$ 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 $J(u+w) \geqslant J(u)$ (tout élément de V peut être écrit sous la forme $u+w$). Nous avons :

$$J(u+w) = \frac{1}{2}\big(a(u,u) + a(u,w) + a(w,u) + a(w,w)\big) - l(u) - l(w) \tag{26.176a}$$

$$= J(u) + \underbrace{a(u,w) - l(w)}_{=0} + \frac{1}{2}a(w,w) \tag{26.176b}$$

$$= J(u) + \frac{1}{2}a(w,w) \tag{26.176c}$$

Vu que a est coercive, le second terme est positif et nous avons

$$J(u+w) \geqslant J(u), \tag{26.177}$$

ce qu'il fallait.

$\Leftarrow$ Soit $u \in V$, un élément minimisant la fonctionnelle J. Nous fixons $w \in V$ et considérons la fonction $g\colon \mathbb{R} \to \mathbb{R}$ définie par

$$g(\epsilon) = J(u + \epsilon w). \tag{26.178}$$

En développant un peu et en regroupant les termes,

$$g(\epsilon) = J(u) + \epsilon\big(a(u,w) - l(w)\big) + \frac{\epsilon^2}{2}a(w,w). \tag{26.179}$$

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 $g(\epsilon) \geqslant g(0)$ ou encore : $\epsilon = 0$ est minimum local (et même global) de g. Le polynôme (26.179) prend son minimum en $\epsilon = 0$ si et seulement si

$$a(u,w) - l(w) = 0. \tag{26.180}$$

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 27

Analyse complexe

27.1 Fonctions holomorphes

La dérivée complexe est discutée à la section 13.26.

27.1.1 Équations de Cauchy-Riemann

Notons que la formule (13.619) 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 $s\colon \mathbb{R} \to \mathbb{C}$ telle que $\lim_{t\to 0} s(t)/t = 0$ et

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + s(|z - z_0|). \tag{27.1}$$

Nous introduisons les opérateurs

$$\frac{\partial}{\partial z} = \partial = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \tag{27.2a}$$

$$\frac{\partial}{\partial \bar{z}} = \bar{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) \tag{27.2b}$$

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 la fonction f a des composantes harmoniques.

Théorème 27.1.
Si $f \in C^1(\Omega)$ alors nous avons équivalence des faits suivants :

(1) f est holomorphe sur Ω,

(2) f vérifie $\partial_{\bar{z}} f = 0$.

Proposition 27.2.
Une application $f\colon \Omega \to \mathbb{C}$ est $\mathbb{C}$-dérivable sur Ω si et seulement si elle est différentiable et

$$\begin{cases} \dfrac{\partial u}{\partial x} = \dfrac{\partial v}{\partial y} & \text{(27.3a)} \\ \dfrac{\partial u}{\partial y} = -\dfrac{\partial v}{\partial x} & \text{(27.3b)} \end{cases}$$

où $f(x + iy) = u(x, y) + iv(x, y)$. Ces équations se notent de façon plus compacte

$$\frac{\partial f}{\partial \bar{z}} = 0. \tag{27.4}$$

Les équations (27.3) sont les équations de **Cauchy-Riemann**.

Démonstration. La différentielle de $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ est donnée par la matrice

$$T = \begin{pmatrix} \partial_x u(a) & \partial_y u(a) \\ \partial_x v(a) & \partial_y v(a) \end{pmatrix}. \tag{27.5}$$

Cette matrice est une similitude si et seulement si les équations de Cauchy-Riemann sont satisfaites. En effet si $1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, la matrice T est une similitude (écrivons $\alpha + i\beta$ son coefficient) si

$$\begin{cases} T(1) = \alpha + i\beta & \text{(27.6a)} \\ T(i) = -\beta + i\alpha, & \text{(27.6b)} \end{cases}$$

c'est-à-dire

$$T = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}. \tag{27.7}$$

Identifier cette matrice à (27.5) fournit le résultat annoncé. □

Proposition 27.3.
Si $f\colon \mathbb{C} \to \mathbb{C}$ est holomorphe, alors nous avons

$$df_{z_0} = (\partial_z f)(z_0) \tag{27.8}$$

au sens où l'opérateur linéaire $df_{z_0}\colon \mathbb{C} \to \mathbb{C}$ est l'opération de multiplication par le nombre complexe $(\partial_z f)(z_0)$.

Démonstration. Soit $f(x+iy) = f_1(x,y) + if_2(x,y)$ une fonction holomorphe. Les fonctions réelles f_1 et f_2 sont assujetties aux équations de Cauchy-Riemann de la proposition 27.2 :

$$\begin{cases} \partial_x f_1 = \partial_y f_2 & \text{(27.9a)} \\ \partial_x f_2 = -\partial_y f_1. & \text{(27.9b)} \end{cases}$$

Nous avons, en recourant à un petit abus de notation entre $f_i\colon \mathbb{R}^2 \to \mathbb{R}$ et $f_i\colon \mathbb{C} \to \mathbb{R}$:

$$\begin{aligned} df_{z_0}(u) &= \frac{d}{dt}\Big[f(z_0 + tu)\Big]_{t=0} && \text{(27.10a)} \\ &= \frac{d}{dt}\Big[f_1(z_0 + tu) + if_2(z_0 + tu)\Big]_{t=0} && \text{(27.10b)} \\ &= \partial_x f_1 u_1 + \partial_y f_1 u_2 + i\big(\partial_x f_2 u_1 + \partial_y f_2 u_2\big) && \text{(27.10c)} \\ &= \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} && \text{(27.10d)} \\ &= \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ -\partial_y f_1 & \partial_y f_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. && \text{(27.10e)} \end{aligned}$$

En utilisant le lemme 13.246 nous reconnaissons la matrice de multiplication par le nombre $\partial_x f_1 - i\partial_y f_1$. Or justement,

$$\partial_z f = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) f = \frac{1}{2}\big(\partial_x f_1 + i\partial_x f_2 - i\partial_y f_1 + \partial_y f_2\big), \tag{27.11}$$

qui se réduit à $\partial_x f_1 - i\partial_y f_1$ lorsque nous y appliquons les équations de Cauchy-Riemann. □

27.1.2 Intégrales sur des chemins fermés

Lemme 27.4.
Si g est une fonction continue dans un ouvert $\Omega \subset \mathbb{C}$ et si g admet une primitive complexe sur Ω alors

$$\int_\gamma g(z)dz = 0 \tag{27.12}$$

pour tout chemin fermé γ de classe C^1 contenu dans Ω.

Démonstration. Nommons G une primitive de g. Par définition,

$$\int_\gamma g = \int_\gamma G' \tag{27.13a}$$
$$= \int_0^1 G'\big(\gamma(t)\big)\gamma'(t)dt \tag{27.13b}$$
$$= \int_0^1 (G\circ g\gamma)'(t)dt \tag{27.13c}$$
$$= G(\gamma(1)) - G\big(\gamma(0)\big) \tag{27.13d}$$
$$= 0 \tag{27.13e}$$

parce que le chemin est fermé : $\gamma(0)=\gamma(1)$. □

Lemme 27.5 (Goursat[395])**.**
Soit Ω un ouvert dans $\mathbb{C}$ et f une fonction continue sur Ω, holomorphe sur Ω moins éventuellement un point (nommé $z_1\in\Omega$). Soit T, un triangle[1] *fermé inclus dans Ω. Alors nous avons*

$$\int_{\partial T} f(z)dz = 0. \tag{27.14}$$

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

$$c = l(\gamma)^{-2}|\int_\gamma f|. \tag{27.15}$$

Notre objectif est de montrer que $c=0$. Soit A,B,C les trois sommes du triangle ; nous divisons le triangle de la façons 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 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) $AB'C'$ de façon à avoir

$$\int_\gamma f = \sum_{i=1}^4 \int_{\partial T_i^{(0)}} f. \tag{27.16}$$

Cela implique que pour au moins un des quatre triangles (disons $T_k^{(0)}$ pour fixer les idées) nous ayons

$$\int_{\partial T_k^{(0)}} f \geqslant \frac{1}{4}\int_{\partial T^{(0)}} f \tag{27.17}$$

Nous notons $T^{(1)}$ ce triangle. Comme noté précédemment nous avons

$$l(\partial T^{(1)}) = \frac{1}{2}l(\partial T^{(0)}), \tag{27.18}$$

et donc

$$l(\gamma_1)^{-2}|\int_{\gamma_1}|f = 4l(\gamma_0)^{-2}|\int_{\gamma_1} f| \geqslant 4l(\gamma_0)^{-2}\frac{1}{4}|\int_{\gamma_0} f| = c. \tag{27.19}$$

En répétant le procédé nous construisons une suite de triangles $T^{(n)}$ qui satisfont toujours

$$l(\partial T^{(n)}) = \frac{1}{2^n}l(\partial T^{(0)}). \tag{27.20}$$

1. Nous considérons ici le triangle « plein ».

Ces triangles forment une suite de fermés emboités dont le diamètre tend vers zéro. Leur intersection contient donc exactement un point (lemme 18.119) que nous nommons z_0 (et qui appartient évidemment à Ω). Étant donné que f est holomorphe nous utilisons le développement limité (27.1) autour de z_0 :

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + s(|z - z_0|)(z - z_0) \tag{27.21}$$

avec $\lim_{t\to 0} s(t) = 0$. Nous posons $g(z) = f(z_0) + f'(z_0)(z - z_0)$ et nous considérons $\epsilon > 0$. Soit $\alpha > 0$ tel que

$$|f(z) - g(z)| < \epsilon|z - z_0| \tag{27.22}$$

pour tout $|z - z_0| < \alpha$. Le α 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

$$|f(z) - g(z)| < \epsilon|z - z_0|. \tag{27.23}$$

Notons que la fonction g est une dérivée : c'est la dérivée de la fonction

$$G(z) = zf(z_0) + \frac{1}{2}f'(z_0)(z - z_0)^2. \tag{27.24}$$

Par conséquent nous avons

$$\int_{\gamma_n} g = 0 \tag{27.25}$$

par le lemme 27.4. Nous avons donc

$$\begin{aligned} |\int_{\gamma_n} f| &= |\int_{\gamma_n} (f - g)| && \text{(27.26a)}\\ &\leqslant l(\gamma_n) \max\{|f(z) - g(z)| \text{ tel que } z \in T^{(n)}\} && \text{(27.26b)}\\ &\leqslant \epsilon l(\gamma_n)^2, && \text{(27.26c)} \end{aligned}$$

et par conséquent

$$c \leqslant l(\gamma_n)^{-2}|\int_{\gamma_n} f| \leqslant \epsilon, \tag{27.27}$$

ce qui signifie que $c = 0$ parce que ϵ 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 ϵ. Le vecteur $v = z_1 - C$ fait par exemple l'affaire. En vertu du point précédent nous avons

$$\int_{\partial T_\epsilon} f = 0 \tag{27.28}$$

pour tout $\epsilon > 0$. Étant donné que la fonction f est continue (y compris en z_1), l'intégrale sur ∂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

$$\int_T f = \int_{Az_1C} f + \int_{ABz_1} f + \int_{BCz_1} f, \tag{27.29}$$

alors que par le point précédent les trois intégrales du membre de droite sont nulles. □

Proposition 27.6 ([395]).
Soient Ω un ouvert étoilé et f une fonction holomorphe sur Ω sauf éventuellement en un point z_1 où f est seulement continue. Alors si γ est un chemin fermé dans Ω, nous avons

$$\int_\gamma f = 0. \tag{27.30}$$

Proposition 27.7.
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 16.25. □

27.1.3 Lacets, indice et homotopie

Proposition-définition 27.8.
Soit γ un chemin fermé[2] *dans $\mathbb{C}$. L'**indice** de la courbe γ est la fonction*

$$\begin{aligned}\mathrm{Ind}_\gamma\colon \mathbb{C}\backslash\gamma &\to \mathbb{Z}\\ z &\mapsto \frac{1}{2\pi i}\int_\gamma \frac{d\omega}{\omega - z}.\end{aligned} \tag{27.31}$$

*Un chemin continu et fermé (au sens $\gamma(1) = \gamma(0)$) est un **lacet**.*

(1) La fonction Ind_γ est continue et prend effectivement des valeurs entières.
(2) La fonction indice est constante sur chaque composante connexe de $\mathbb{C}\backslash\gamma$ et est nulle sur la composante non bornée.

Le second point est en partie la proposition 7.39.

Définition 27.9.
*Si γ_1 et γ_2 sont deux lacets en $x_0 \in X$ (un espace topologique), une **équivalence d'homotopie** est une application $f\colon [0,1]\times[0,1] \to X$ telle que*

(1) $f(0,t) = \gamma_1(t)$ pour tout t ;
(2) $f(1,t) = \gamma_1(t)$ 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.

Exemple 27.10
Si γ est un cercle de centre $z_0 \in \mathbb{C}$ et de rayon r, alors

$$\mathrm{Ind}_\gamma(z) = \begin{cases} 2\pi i & \text{si } z \in B(z_0, r)\\ 0 & \text{sinon.}\end{cases} \tag{27.32}$$

La seconde ligne provient directement de la proposition 27.8. Pour la première, le cercle γ se paramètre par

$$\gamma(\theta) = z_0 + re^{i\theta}, \tag{27.33}$$

et l'intégrale vaut

$$\int_\gamma \frac{d\omega}{\omega - z_0} = \int_0^{2\pi} \frac{1}{re^{i\theta}} ire^{i\theta} d\theta = 2\pi i. \tag{27.34}$$

L'indice de ce chemin va évidemment jouer un rôle particulier dans la suite. △

Théorème 27.11 (Cauchy, version homotopique[396])**.**
Soient Ω un ouvert de $\mathbb{C}$ et f une fonction holomorphe sur Ω. Si γ_1 et γ_2 sont deux lacets homotopes de classe C^1 dans Ω, alors

$$\int_{\gamma_1} f(z)dz = \int_{\gamma_2} f(z)dz. \tag{27.35}$$

Corollaire 27.12 ([396])**.**
Soient $a \in \mathbb{C}$ ainsi que deux chemins γ_1 et γ_2 homotopes dans $\mathbb{C}\backslash\{a\}$. Alors $\mathrm{Int}(\gamma_1, a) = \mathrm{Ind}(\gamma_2, a)$.

Il y a aussi des choses sur l'indice dans [395].

2. Par abus de langage, nous désignerons par γ à la fois le chemin et son image.

27.1.4 Théorème de Cauchy et analycité

Cette sous-section veut prouver le théorème de Cauchy. Comme d'habitude, une référence qui ne peut pas rater est [395].

Théorème 27.13 (Formule de Cauchy)**.**
Soient Ω ouvert dans $\mathbb{C}$, $z_0 \in \Omega$ et f une fonction holomorphe sur Ω. Soit $r > 0$ tel que $B(z_0, r) \subset \Omega$. Alors pour tout $z \in B(z_0, r)$ nous avons

$$f(z) = \frac{1}{2\pi i} \int_{\partial B(z_0,r)} \frac{f(\omega)}{\omega - z} d\omega. \tag{27.36}$$

Démonstration. Soit $z \in B(z_0, r)$. Considérons la fonction

$$g(\omega) = \begin{cases} \frac{f(\omega)-f(z)}{\omega - z} & \text{si } \omega \neq z \\ f'(z) & \text{si } \omega = z. \end{cases} \tag{27.37}$$

Cette fonction est holomorphe sur $B(z_0, r)\backslash\{z\}$ et continue en z. Elle vérifie donc la proposition 27.6 et nous avons

$$\int_\gamma g = 0 \tag{27.38}$$

où γ est le cercle de centre z_0 et de rayon r. Nous avons donc

$$0 = \int_\gamma \frac{f(\omega)}{\omega - z} - \int_\gamma \frac{f(z)}{\omega - z}, \tag{27.39}$$

et ayant déjà calculé la seconde intégrale dans l'exemple 27.10 nous en déduisons

$$\int_\gamma \frac{f(\omega)}{\omega - z} d\omega = 2\pi i f(z), \tag{27.40}$$

ce qu'il fallait. □

Définition 27.14.
*Une fonction $f : \Omega \to \mathbb{C}$ est $\mathbb{C}$-**analytique** sur Ω si pour tout $z_0 \in \Omega$, il existe une suite complexe (c_n) et $r > 0$ tels que*

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n \tag{27.41}$$

pour tout $z \in B(z_0, r)$.

Théorème 27.15.
Soient Ω ouvert dans $\mathbb{C}$ et f holomorphe sur Ω. Soient encore $z_0 \in \Omega$ et r_0 tels que $B(z_0, r_0) \subset \Omega$. Alors :

(1) Sur $B(z_0, r_0)$, la fonction f s'écrit

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \tag{27.42}$$

(2) Nous avons

$$a_n = \frac{f^{(n)}(z_0)}{n!} = \frac{1}{2\pi i} \int_\gamma \frac{f(\omega)}{(\omega - z_0)^{n+1}} d\omega \tag{27.43}$$

où $\gamma = \partial B(z_0, r)$ avec $|z - z_0| < r < r_0$.

(3) En particulier f est infiniment dérivable.

Démonstration. Soit $r > 0$ tel que $|z - z_0| < r < r_0$. La formule de Cauchy (théorème 27.13) nous dit que

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(\omega)}{\omega - z}d\omega \tag{27.44}$$

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

$$f(z) = \frac{1}{2\pi i}\int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{z_0 + re^{i\theta} - z} rie^{i\theta}d\theta \tag{27.45a}$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{1 - e^{-i\theta}(z - z_0)/r}d\theta. \tag{27.45b}$$

Nous pouvons développer l'intégrante en puissance de $(z - z_0)$ en utilisant la formule 16.359. Ici le rôle de x est tenu par

$$e^{-i\theta}(z - z_0)/r \tag{27.46}$$

dont le module est bien plus petit que 1, par hypothèse sur r. Nous avons donc

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \sum_{n=0}^{\infty} f(z_0 + re^{i\theta})e^{-in\theta}r^{-n}(z - z_0)^n d\theta. \tag{27.47}$$

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

$$M\left|\frac{z - z_0}{r}\right|^n \tag{27.48}$$

où M est le maximum de $|f|$ sur γ. La borne (27.48) ne dépend pas de θ ; par conséquent la convergence de la somme est uniforme en θ par le critère de Weierstrass (théorème 13.295). Le théorème 16.1 s'applique[3] et nous pouvons permuter la somme avec l'intégrale.

Ce que nous trouvons est que

$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n \tag{27.49}$$

où

$$a_n = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{i\theta})e^{-in\theta}r^{-n}d\theta = \frac{1}{2\pi i}\int_\gamma \frac{f(\omega)}{(\omega - z_0)^{n+1}}. \tag{27.50}$$

Cette formule est valable pour $|z - z_0| < r$. Sur cette boule, la fonction est donc une série entière. Le théorème de Taylor 16.100 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

$$a_n = \frac{f^{(n)}(z_0)}{n!}. \tag{27.51}$$

□

Corollaire 27.16.
Soit f une fonction continue sur un ouvert Ω telle que pour toute boule $B(a, r)$ contenue dans Ω, nous ayons

$$f(a) = \frac{1}{2\pi i}\int_{\partial B(a,r)} \frac{f(\xi)}{\xi - a}d\xi. \tag{27.52}$$

Alors f est holomorphe.

3. Étant donné que nous savions déjà que la somme était une fonction intégrable, nous sommes loin d'avoir utilisé toute la puissance du théorème.

Démonstration. Il suffit de recopier la démonstration du théorème 27.15 pour savoir que f se développe en série de puissances et est donc en particulier dérivable. □

Le fait qu'une fonction holomorphe soit C^∞ comme dit dans la proposition 27.15 permet de démonter un résultat de dérivation sous l'intégrale, qui dépend de pouvoir majorer la différentielle.

Proposition 27.17.
Soit une fonction continue $g: \mathbb{R} \times \mathbb{C} \to \mathbb{C}$. Nous supposons que pour tout t, la fonction $z \mapsto g(t,z)$ est $\mathbb{C}$-dérivable (définition 13.243) et différentiable. Soit B compact dans $\mathbb{R}$ et la fonction

$$G(z) = \int_B g(t,z)dt. \tag{27.53}$$

que nous supposons exister pour tout z.

Alors

$$G'(z) = \int_B g'(t,z)dt \tag{27.54}$$

où le prime réfère à la $\mathbb{C}$-dérivée par rapport à la variable z à t fixé.

Démonstration. Nous fixons $z \in \mathbb{C}$ et nous considérons la suite de fonctions

$$g_i(t) = \frac{g(t, z+\epsilon_i) - g(t,z)}{\epsilon_i} \tag{27.55}$$

où ϵ_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 $\lim_{i\to\infty} g_i(t) = g'(t,z)$. Et vu que g est supposée dérivable, c'est le cas.

Nous avons aussi, par linéarité de l'intégrale :

$$G'(z) = \lim_{i\to\infty} \int_B g_i(t)dt. \tag{27.56}$$

La difficulté est de permuter la limite et l'intégrale. Pour cela nous allons utiliser la convergence dominée de Lebesgue (théorème 15.184). Afin de majorer $|g_i(t)|$ par une fonction intégrable en t (uniformément en i), nous exploitons le théorème des accroissements finis, théorème 12.151. En notant dg la différentielle de g par rapport à z à t fixé, pour chaque t et chaque i nous avons

$$|g(t, z+\epsilon_i) - g(t,z)| \leqslant \sup_{\xi\in[z,z+\epsilon_i]} \|dg_\xi\|\|\epsilon_i\|. \tag{27.57}$$

Vu que z est fixé et que ξ 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'analycité par le théorème 27.15), nous pouvons majorer $\|dg_\xi\|$ par une constante $M_i(z)$ qui dépend a 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 $|\epsilon_i| \to 0$. 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

$$|g(t, z+\epsilon_i) - g(t,z)| \leqslant M(z)\|\epsilon_i\| \tag{27.58}$$

et donc $|g_i(t)| \leqslant M(z)$. La constante (par rapport à t) $M(z)$ est évidemment intégrable sur le compact B et nous pouvons permuter la limite avec l'intégrale :

$$G'(z) = \lim_{i\to\infty} \int_B g_i(t)dt = \int_B \lim_{i\to\infty} g_i(t)dt = \int_B g'(t,z)dt. \tag{27.59}$$

□

Proposition 27.18.
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 13.625.

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

$$\lim_{h\to 0}\left|\frac{F(z+h)-F(z)-dF_z(h)}{h}\right| = 0. \tag{27.60}$$

La limite vaut évidemment encore zéro si nous enlevons les modules :

$$0 = \lim_{h\to 0}\frac{F(z+h)-F(z)-f(z)h}{h} \tag{27.61a}$$

$$= \lim_{h\to 0}\frac{F(z+h)-F(z)}{h} - f(z). \tag{27.61b}$$

Donc $F' = f$. Cela montre que F est $\mathbb{C}$-dérivable et donc holomorphe. En conséquence du théorème 27.15, la fonction F est infiniment dérivable et f l'est alors aussi. La fonction f est donc holomorphe [4]. □

27.1.5 Théorème de Brouwer en dimension 2

Pour d'autres versions du théorème de Brouwer, voir la sous-section 21.5.1.

Théorème 27.19 (Brouwer en dimension 2[43])**.**
Soit $\mathcal{B}$ 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 21.30 nous savons que la fonction

$$\begin{aligned} g\colon \overline{B(0,1)} &\to \partial B(0,1)\\ x &\mapsto \lambda(x)\big(x-f(x)\big)+f(x) \end{aligned} \tag{27.62}$$

est continue. De plus $g(x) = x$ sur $\partial B(0,1)$. Nous allons montrer qu'une telle fonction [5] ne peut pas exister.

Pour $s \in [0,1]$ nous paramétrons le cercle $\partial B(0,s)$ par

$$\begin{aligned} x_s\colon [0,1] &\to \partial B(0,1)\\ t &\mapsto \big(s\cos(2\pi t), s\sin(2\pi t)\big). \end{aligned} \tag{27.63}$$

Ensuite nous considérons les chemins

$$\begin{aligned} \gamma_s\colon [0,1] &\to \partial B(0,s)\\ t &\mapsto g\circ x_s. \end{aligned} \tag{27.64}$$

L'application γ_s est continue et $\gamma_s(0) = \gamma_s(1)$. Les chemins γ_s sont des lacets ; nous nous intéressons maintenant à l'indice au point 0 de γ_0 et γ_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 27.8,

$$\mathrm{Ind}_{\gamma_0}(0) = \frac{1}{2\pi i}\mathrm{Ind}_{\gamma_0}\frac{d\omega}{\omega} = \frac{1}{2\pi i}\int_0^1 \frac{\gamma_0'(t)}{\gamma_0(t)}dt = 0, \tag{27.65}$$

4. Dire que la dérivée d'une fonction holomorphe est holomorphe est un raisonnement classique.
5. Qui est nommée *rétraction* de la sphère sur elle-même.

alors que

$$\mathrm{Ind}_{\gamma_1}(0) = \frac{1}{2\pi i}\int_0^1 \frac{2i\pi e^{2i\pi t}}{e^{2i\pi t}}dt = 1. \tag{27.66}$$

Nous considérons l'homotopie

$$\begin{aligned}\gamma\colon [0,1]\times[0,1] &\to \overline{B(0,1)}\\ (s,t) &\mapsto \gamma_s(t) = (g\circ x_s)(t).\end{aligned} \tag{27.67}$$

Nous avons $g(0) \neq 0$ parce que g prend ses valeurs sur le bord. Vu que c'est une équivalence d'homotopie[6] entre γ_1 et γ_2, les indices devraient être égaux par le corollaire 27.12. □

27.1.6 Principe des zéros isolés

Théorème 27.20 (Principe des zéros isolés [395])**.**
Soient f une fonction holomorphe et a, une 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[7] autour de a :

$$f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n \tag{27.68}$$

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

$$f(z) = c_m(z-a)^m\big(1+\sum_{k=1}^{\infty} d_k(z-a)^k\big) \tag{27.69}$$

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 (27.68) parce que la suite $\frac{d_k r^{m+k}}{c_m}$ reste bornée (critère d'Abel, lemme 16.10). Si nous posons

$$g(z) = 1+\sum_{k=1}^{\infty} d_k(z-a)^k, \tag{27.70}$$

alors g est une fonction continue et $g(a) = 1$. De plus

$$f(z) = c_m(z-a)^m g(z). \tag{27.71}$$

Soit une suite (z_n) de zéros de f qui converge vers a. Étant donné que g est continue, nous devrions avoir $\lim_{k\to\infty} g(z_k) = g(a) = 1$, 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. □

Corollaire 27.21.
Soit f une fonction holomorphe sur un ouvert connexe Ω. Si f s'annule sur un ouvert (non vide) de Ω, alors f s'annule sur tout Ω.

Démonstration. soit

$$N = \{z \in \Omega \text{ tel que } f = 0 \text{ sur un ouvert autour de } z\}. \tag{27.72}$$

6. Définition 27.9
7. Possible par le théorème 27.15 parce que f est holomorphe.

Le fait que N soit ouvert est évident à partir de sa définition. Nous allons montrer que N est également fermé dans Ω, 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

$$f(z) = \lim_{n\to\infty} f(z_n) = 0, \tag{27.73}$$

ce qui fait de z un zéro non isolé de f. Par conséquent le principe des zéros isolés (théorème 27.20) 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. □

Définition 27.22.
*Si K est un compact de $\mathbb{C}$, nous disons qu'une fonction est **holomorphe** sur K si il existe un ouvert contenant K sur lequel f est holomorphe.*

Et si f n'est réellement définie que sur K, elle est holomorphe sur K si il existe une extension holomorphe de f vers un ouvert contenant K.

Lemme 27.23.
Si f est une fonction holomorphe sur le compact K, alors il existe une fonction 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.

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[8] de f dans K. Vu qu'une suite dans un compact contient une sous-suite convergente (théorème 7.97), la fonction f aurait un point d'accumulation de racines. Alors le principe des zéros isolés (théorème 27.20) nous donne un ouvert sur lequel f est nulle et donc le corollaire 27.21 nous dit que f est identiquement nulle.

Autour d'une racine Bref, la fonction f possède un nombre fini de racines sur K. Soit z_0 l'un d'eux.

Par le théorème 27.15(1), nous avons, sur un voisinage de z_0 :

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \tag{27.74}$$

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

$$f(z) = (z - z_0)^k g(z) \tag{27.75}$$

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 27.7 nous dit que g est $\mathbb{C}$-dérivable. C'est-à-dire holomorphe par définition.

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 :

$$h(z) = \begin{cases} \dfrac{f(z)}{\prod_k P_k(z)} & \text{si } z \neq z_i \\ \dfrac{g_i}{\prod_{k\neq i} P_k(z)} & \text{si } z = z_i. \end{cases} \tag{27.76}$$

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 13.251.

8. Notez l'utilisation de la proposition 1.26 que je vous invite à ne pas considérer comme une trivialité absolue.

Pour les autres notons que pour tout $z \in B(z_i, r_i)$,

$$h = \frac{g_i}{\prod_{k \neq i} P_k}. \tag{27.77}$$

Cela est encore un quotient dont le dénominateur ne s'annule pas [9].

La réponse Nous avons, pour tout $z \in K$:

$$f(z) = h(z) \prod_k P_k(z). \tag{27.78}$$

□

Afin de détendre l'atmosphère, nous allons laisser tomber l'analyse quelques instants et prouver un résultat d'algèbre.

Proposition 27.24 ([1, 397]).
L'anneau des fonctions holomorphes sur un compact [10] donné de $\mathbb{C}$ est principal [11].

Démonstration. Nous nommons A l'ensemble des fonctions holomorphes sur le compact K, et J un anneau de A.

A est un anneau Le point délicat de la définition 1.37 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 holomorphes » sont valides sur des ouverts alors que nous sommes ici sur un compact. Soient f et g dans A ; nous nommons Ω_f et Ω_g des ouverts contenant K tels que f est holomorphe sur Ω_f et g sur Ω_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 holomophes.

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 27.23. 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

$$S = \{P_f \text{ tel que } f \in J\} \tag{27.79}$$

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 $(S) = J$.

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 3.163 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

$$M = \mathbb{C}[X]p. \tag{27.80}$$

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 « $M = (p)$ ».

$A\mathbb{C}[X] = A$ L'inclusion $A \subset A\mathbb{C}[X]$ est dûe 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.

Suite des opérations Nous avons :

$$J = AS \subset A\mathbb{C}[X]p. \tag{27.81}$$

9. Nous avons choisi les r_i de telle sorte que les boules ne s'intersectent pas.
10. Être holomorphe sur un compact signifie qu'il existe une extension holomorphe à un ouvert contenant le compact.
11. Définition 3.96

Voilà une inclusion de montrée. Reste à faire l'autre.
Vu que $p \in J$ nous avons aussi $Ap \subset J$. Et donc

$$A\mathbb{C}[X]p = Ap \subset J. \tag{27.82}$$

Avec ces deux inclusions, $J = A\mathbb{C}[X]p = Ap$. Donc J est engendré par un seul élément et est principal.

□

27.1.7 Prolongement de fonctions holomorphes

Proposition 27.25.
Soient Ω un ouvert de $\mathbb{C}$ et $f : \Omega \to \mathbb{C}$ une fonction holomorphe sur $\Omega\backslash\{a\}$ ($a \in \Omega$). 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 Ω.

Le théorème de prolongement de Riemann 27.39 donnera plus d'informations.

Démonstration. Nous définissons la fonction $g : \Omega \to \mathbb{C}$ par

$$g(z) = \begin{cases} (z-a)f(z) & \text{si } z \neq a \\ 0 & \text{si } z = a. \end{cases} \tag{27.83}$$

Sur $\Omega\backslash\{a\}$, la fonction g est holomorphe (produit de fonctions holomorphes), et elle est continue en a. Par conséquent elle est holomorphe sur Ω. Nous la développons en série entière sur une boule $B(a, r)$:

$$g(z) = \sum_{n=0}^{\infty} c_n (z-a)^n. \tag{27.84}$$

Nous avons $g(a) = c_0 = 0$. Nous posons

$$\varphi(z) = \sum_{n=0}^{\infty} c_{n+1} (z-a)^n. \tag{27.85}$$

Si $z \neq a$, alors $\varphi(z) = f(a)$ parce que $\varphi(z) = g(z)/(z-a)$. Mais φ est continue en a, et donc holomorphe en a.

La fonction φ est par conséquent un prolongement holomorphe de f en a. □

27.1.8 Théorème de Runge

Le théorème que nous allons prouver n'est en réalité qu'une partie de ce qui est usuellement appelle le théorème de Runge.

Théorème 27.26 (Théorème de Runge)**.**
Soit K, un compact de $\mathbb{C}$ tel que $\complement K$ soit connexe. Si $a \in \complement K$ alors la fonction

$$\varphi_a(z) = \frac{1}{z-a} \tag{27.86}$$

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 $a \in \complement K$, la fonction φ_a est dans $P(K)$. Pour cela nous considérons l'ensemble

$$A = \{a \in \complement K \text{ tel que } \varphi_a \in P(K)\} \tag{27.87}$$

et nous allons montrer qu'il est à la fois non vide, ouvert et fermé dans le connexe $\complement K$.

Je répète : nous allons prouver l'ouverture et la fermeture *pour la topologie de* $\complement K$. Nous n'allons pas prouver que A est un ouvert de $\mathbb{C}$. Ce qui sera par conséquent prouvé est que $A = \complement K$.

Non vide Soit $R = \sup_{z\in K} |z|$ et $a \in \complement K$ tel que $|a| > R$. Nous avons

$$\varphi_a(z) = \frac{1}{a}\frac{1}{\frac{z}{a}-1} = -\frac{1}{a}\frac{1}{1-\frac{z}{a}} = -\frac{1}{a}\sum_{k=0}^{\infty}\left(\frac{z}{a}\right)^k = \sum_{k=0}^{\infty}\frac{z^k}{a^{k+1}}. \tag{27.88}$$

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

$$P_n(z) = \sum_{k=0}^{n}\frac{z^k}{a^{k+1}} \tag{27.89}$$

converge uniformément sur $B(0,R)$ et en particulier sur K. Donc $P_n \to \varphi_a$.

Fermé Nous allons montrer que la fermeture de A (dans $\complement K$) est inclue dans A, et donc qu'elle est égale à A et donc que A est fermé. Par le lemme 7.12, la fermeture de A dans $\complement K$ est l'ensemble $\bar{A} \cap \complement K$ où $\bar{A}$ est la fermeture de A au sens usuel.

Bref, soit $a \in \bar{A} \cap \complement K$, 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 $\bar{A} \cap \complement A = A$ 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 a 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

$$|\varphi_{a_n}(z) - \varphi_a(z)| = \left|\frac{1}{z-a_n} - \frac{1}{z-a}\right| = \left|\frac{a_n - a}{(z-a_n)(z-a)}\right|. \tag{27.90}$$

Vu que $a_n \in B(a, \frac{d}{2})$ et que $z \in K$ et $d = d(a,K)$ nous avons $|a_n - z| \geqslant \frac{d}{2}$; et aussi $|a - z| \geqslant \frac{d}{2}$. Nous pouvons donc majorer (27.90) par

$$|\varphi_{a_n}(z) - \varphi_a(z)| \leqslant 2\frac{|a_n - a|}{d^2}. \tag{27.91}$$

Donc nous avons

$$\|\varphi_a - \varphi_{a_n}\|_K \leqslant 2\frac{|a_n - a|}{d^2} \to 0 \tag{27.92}$$

où la norme $\|.\|_K$ est la norme supremum sur K. Donc $a \in \overline{P(K)} = P(K)$ et A est fermé.

Ouvert Vu que K est compact, il est fermé et donc $\complement K$ est ouvert. Par conséquent, ainsi que précisé dans l'exemple 7.13, les ouverts de $\complement K$ sont les ouverts de $\mathbb{C}$ contenus dans $\complement K$. 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

$$\varphi_{a+h}(z) = \frac{1}{z-a-h} = \frac{1}{z-a}\frac{1}{1-\frac{h}{z-a}} = \sum_{k=0}^{\infty}\frac{h^k}{(z-a)^{k+1}}. \tag{27.93}$$

Déjà ici nous demandons $h < \sup_{z\in K}|z-a|$. Puisque $|z-a| > d$, nous avons alors

$$|\varphi_{a+h}(z)| \leqslant \sum_{k=0}^{\infty}\frac{h^k}{d^{k+1}} < \infty. \tag{27.94}$$

Cela pour dire que la somme à droite de (27.93) converge bien pourvu que h soit bien petit. Nous pouvons donc poursuivre :

$$\varphi_{a+h}(z) = \sum_{k=0}^{\infty}\frac{h^k}{(z-a)^{k+1}} = \sum_{k=0}^{\infty}h^k\varphi_a(z)^{k+1}. \tag{27.95}$$

Nous montrons maintenant que la convergence de la somme (27.95) est en réalité uniforme en z. En effet

$$\left|\varphi_{a+h}(z) - \sum_{k=0}^{N} h^k \varphi_a(z)^{k+1}\right| = \left|\sum_{k=N+1}^{\infty} h^k \varphi_a(z)^{k+1}\right| \tag{27.96a}$$

$$\leqslant \sum_{k=N+1}^{\infty} |h|^k |\varphi_a(z)|^{k+1}. \tag{27.96b}$$

Étant donné que φ_a est continue sur le compact K, elle y est majorée en module ; on peut même être plus précis :

$$|\varphi_a(z)| = \frac{1}{|z-a|} \leqslant \frac{1}{d}. \tag{27.97}$$

Nous pouvons donc écrire

$$\left|\varphi_{a+h}(z) - \sum_{k=0}^{N} h^k \varphi_a(z)^{k+1}\right| \leqslant \frac{1}{d} \sum_{k=N+1}^{\infty} \left|\frac{h}{d}\right|^k. \tag{27.98}$$

Étant donné que la somme $\sum_{k=0}^{\infty} |h/d|^k$ converge, la limite $N \to \infty$ est nulle et nous avons

$$\lim_{N\to\infty} \|\varphi_{a+h} - \sum_{k=0}^{N} h^k \varphi_a^{k+1}\|_K = 0. \tag{27.99}$$

Pour avoir $\varphi_{a+h} \in P(K)$, il faut encore savoir si les fonctions φ_a^k sont dans $P(K)$ pour tout k. Dans ce cas pour chaque N la somme sera encore dans $P(K)$ et φ_{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 φ_a. Nous allons montrer qu'alors la suite de polynômes P_n^k converge uniformément vers φ_a^k. Soit n tel que $\|P_n - \varphi_a\|_K \leqslant \epsilon$ et utilisons le produit remarquable

$$a^k - b^k = (a-b) \sum_{i=0}^{k-1} a^i b^{k-1-i} \tag{27.100}$$

pour obtenir

$$|P_n(z)^k - \varphi_a(z)^k| \leqslant |P_n(z) - \varphi_a(z)| \sum_{i=0}^{k-1} |P_n(z)^i \varphi_a(z)^{k-1-i}|. \tag{27.101}$$

Vu que P_n et φ_a sont continues sur le compact K, on peut majorer la somme par une constante M, et il restera

$$|P_n(z)^k - \varphi_a(z)^k| \leqslant M |P_n(z) - \varphi_a(z)|, \tag{27.102}$$

ou encore

$$\|P_n^k - \varphi_a^k\| \leqslant M\epsilon. \tag{27.103}$$

Cela prouve que $\varphi_a^k \in P(K)$ et donc que φ_{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 $\complement K$.

Conclusion L'ensemble A est non vide, ouvert et fermé dans $\complement K$, donc il est égal à $\complement K$. Le théorème est ainsi démontré.

□

27.2 Intégrales de fonctions holomorphes

Nous commençons par le lemme technique.

Lemme 27.27 ([395]).
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

$$|f'(z)| \leqslant \frac{r}{\big(r - |z - z_0|\big)^2} \max\big\{f(z_0 + re^{i\theta})\big\}_{\theta\in\mathbb{R}}. \tag{27.104}$$

Démonstration. Par translation nous pouvons supposer que $z_0 = 0$. Étant donné que f est holomorphe, elle admet un développement en séries entières

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \tag{27.105}$$

et nous notons $M = \max\{f(z) \text{ tel que } z \in \overline{B(0, r)}\}$. Nous avons $r^n|a_n| \leqslant M$. Par conséquent

$$\begin{aligned}
|f'(z)| &= \left|\sum_{n=1}^{\infty} n a_n z^{n-1}\right| && \text{(27.106a)}\\
&\leqslant \frac{1}{r}\sum r^n |a_n| n \left(\frac{|z|}{r}\right)^{n-1} && \text{(27.106b)}\\
&< \frac{M}{r}\sum n\left(\frac{|z|}{r}\right)^{n-1}. && \text{(27.106c)}
\end{aligned}$$

À ce point nous devons utiliser la série de l'exemple 16.108. Nous avons alors

$$|f'(z)| \leqslant \frac{M}{r}\frac{1}{\Big(1 - \frac{|z|}{r}\Big)^2} = \frac{Mr}{(r - |z|)^2}. \tag{27.107}$$

□

Théorème 27.28 (Holomorphie sous l'intégrale[395]).
Soit un espace mesuré (Ω, μ), un ouvert A dans $\mathbb{C}$ et une fonction $f\colon A \times \Omega \to \mathbb{C}$. Nous voulons étudier la fonction

$$F(z) = \int_\Omega f(z, \omega) d\mu(\omega) \tag{27.108}$$

pour tout $z \in A$. Nous supposons que

1. *la fonction $f(., \omega)$ est holomorphe sur A pour chaque ω.*
2. *La fonction $f(z, .)$ est mesurable sur (Ω, μ).*
3. *Pour tout compact $K \subset A$, il existe une fonction $g_K\colon \Omega \to \mathbb{R}$ telle que $|f(z, \omega)| \leqslant g_K(\omega)$ et telle que*
$$\int_\Omega g_K(\omega) d\mu(\omega) \tag{27.109}$$
existe.

Alors la fonction F est holomorphe et

$$F'(z) = \int_\Omega \frac{\partial f}{\partial z}(z, \omega) d\mu(\omega). \tag{27.110}$$

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

$$\begin{aligned}
f_\omega\colon \overline{B(z_0, r)} &\to \mathbb{C}\\
z &\mapsto f(z, \omega).
\end{aligned} \tag{27.111}$$

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

$$|f_\omega(z)|\leqslant f_K(\omega) \tag{27.112}$$

pour tout $z\in S$. En utilisant la lemme 27.27 nous avons

$$|f'_\omega(z)|\leqslant \frac{r}{(r-|z-z_0|)^2}\max\{f(z_0+re^{i\theta})\}_{\theta\in\mathbb{R}} \tag{27.113a}$$

$$\leqslant \frac{rf_K(\omega)}{(r-|z-z_0|)^2}. \tag{27.113b}$$

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

$$|f'(z)|\leqslant \frac{rf_K(\omega)}{\left(r-\frac{r}{2}\right)^2}=\frac{4}{r}f_K(\omega). \tag{27.114}$$

Étant donné que la boule $B(z_0,r/2)$ est convexe, la fonction f_ω est Lipschitz et pour tout $h\in\mathbb{C}$ tel que $|h|<r/2$ nous avons

$$\left|\frac{f_\omega(z_0+h)-f_\omega(z_0)}{h}\right|\leqslant\frac{4f_K(\omega)}{r}. \tag{27.115}$$

Soit maintenant une suite (h_n) qui converge vers 0 dans $\mathbb{C}$. Nous considérons la suite de fonctions correspondantes

$$g_n(\omega)=\frac{f(z_0+h_n,\omega)-f(z_0,\omega)}{h_n}. \tag{27.116}$$

Cette suite de fonctions vérifie la convergence ponctuelle

$$g_n(\omega)\to\frac{\partial f}{\partial z}(z_0,\omega). \tag{27.117}$$

De plus g_n est une fonction (de ω) dominée par $\frac{4f_K}{r}$ qui est intégrable. Par conséquent le théorème de la convergence dominée nous indique que

$$\int_\Omega g_n(\omega)d\mu(\omega)\to\int_\Omega\frac{\partial f}{\partial z}(z_0,\omega)d\mu(\omega), \tag{27.118}$$

tandis que

$$F'(z)=\lim_{n\to\infty}\frac{F(z_0+h_n)-F(z_0)}{h_n}=\lim_{n\to\infty}\int_\Omega g_N(\omega)d\mu(\omega). \tag{27.119}$$

□

Corollaire 27.29.
Si f est une fonction holomorphe sur l'ouvert ξ 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 pas la formule suivante :

$$f^{(k)}(z)=\frac{1}{2\pi i}\int_{\partial B(z_0,r)}\frac{f(\xi)}{(\xi-z)^{k+1}}d\xi. \tag{27.120}$$

Démonstration. Nous faisons par récurrence.

Pour la dérivée première Nous appliquons le théorème 27.28 à la fonction

$$g(z,\xi)=\frac{f(\xi)}{\xi-z} \tag{27.121}$$

avec $\xi = \partial B(z_0, r)$ et $A = B(z_0, \rho)$ avec $\rho < r$. Étant donné que f est holomorphe, elle 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

$$|g(z, \xi)| \leqslant \frac{M}{r - \rho}. \tag{27.122}$$

La fonction constante $g_K = \frac{M}{r-\rho}$ est évidemment intégrable. Le théorème conclut que f est holomorphe (cela, nous le savions déjà [12]), et

$$f'(z) = \frac{1}{2i\pi} \int_{\partial B} \frac{f(\xi)}{(\xi - z)^2} d\xi. \tag{27.123}$$

Les dérivées suivantes Pour la récurrence[398] nous supposons que

$$f^{(k)}(z) = \frac{k!}{2i\pi} \int_{\partial B(z_0,r)} \frac{f(\xi)}{(\xi - z)^{k+1}} d\xi, \tag{27.124}$$

et nous tentons de calculer $f^{(k+1)}(z)$. Pour cela nous paramétrons l'intégrale de façon très usuelle :

$$f^{(k)}(z) = \frac{k!}{2\pi i} \int_0^{2\pi} \frac{f(re^{it})}{(re^{it} - z)} ire^{it} dt. \tag{27.125}$$

Nous permettons de permuter la $\mathbb{C}$-dérivation (par rapport à z) et l'intégrale en vertu de la proposition 27.17 appliquée à la fonction

$$g(t, z) = \frac{f(re^{it})}{(re^{it} - z)^k + 1} ire^{it}. \tag{27.126}$$

Cela donne

$$f^{(k+1)}(z) = \frac{k!}{2i\pi} \int_0^{2\pi} f(re^{it}) ire^{it} \frac{k+1}{(re^{it} - z)^{k+1}} dt = \frac{(k+1)!}{2i\pi} \int_{\partial B} \frac{f(\xi)}{(\xi - z)^{k+2}} d\xi. \tag{27.127}$$

□

Théorème 27.30.
Si f est une fonction holomorphe sur le disque ouvert $B(z_0, R)$ alors

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n \tag{27.128}$$

et cette série converge uniformément sur tout compact.

Démonstration. Sans perte de généralité nous supposons que $z_0 = 0$. La formule de Cauchy (théorème 27.13) fournit, pour $z \in B(0, R)$,

$$f(z) = \frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{\xi - z} d\xi = \frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{1 - (z/\xi)} \frac{d\xi}{\xi}. \tag{27.129}$$

En particulier notons que $z \in B(0, R)$ alors que ξ est sur le bord de cette boule ouverte. Donc $|\xi| > |z|$ pour tous les ξ et z qui interviennent. Nous utilisons la série géométrique

$$\frac{1}{1 - (z/\xi)} = \sum_{n=0}^{\infty} \left(\frac{z}{\xi}\right)^n. \tag{27.130}$$

12. Et cela fournit une preuve alternative à la réciproque du théorème de Cauchy : une fonction continue qui vérifie la formule de Cauchy est holomorphe.

Permuter une intégrale et une somme En utilisant la mesure de comptage[13] sur $\mathbb{N}$ (qui est σ-finie), nous pouvons écrire

$$\int_{\partial B} \sum_{n=0}^{\infty} \frac{z^n f(\xi)}{\xi^{n+1}} d\xi = \int_{\partial B} \left(\int_{\mathbb{N}} g(\xi, n) dm(n) \right) d\xi \tag{27.131}$$

où

$$\begin{aligned} g\colon \partial B \times \mathbb{N} &\to \mathbb{C} \\ (\xi, n) &\mapsto \frac{z^n f(\xi)}{\xi^{n+1}}. \end{aligned} \tag{27.132}$$

Nous allons permuter les intégrales en utilisant le théorème de Fubini, selon la procédure décrite en 15.260. Nous commençons par l'intégrale sur $\mathbb{N}$:

$$\int_{\mathbb{N}} |g(n, \xi)| = |\frac{f(\xi)}{\xi}| \sum_{n \in \mathbb{N}} |\frac{z}{\xi}|^n = \frac{1}{R} |f(\xi)| \frac{1}{1 - |z|/R}. \tag{27.133}$$

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 ∂B. Cela parce que B est une boule fermée dans l'ouvert Ω sur lequel f est continue. Au final l'expression à droite de (27.133) est bornée sur le compact ∂B et son intégrale donne un nombre fini.

Tout ceci pour invoquer le corollaire 15.258 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 15.259 pour permuter les intégrales.

Une fois la somme et l'intégrale permutées, nous avons

$$f(z) = \frac{1}{2\pi i} \sum_{n=0}^{\infty} \int_{\partial B} \frac{z^n f(\xi)}{\xi^{n+1}} = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_{\partial B} \frac{f(\xi)}{\xi^{n+1}} \right) z^n. \tag{27.134}$$

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 (27.120).

□

Proposition 27.31 (Morera [399])**.**
Soit Ω ouvert dans $\mathbb{C}$ et f continue. Si

$$\int_{\partial T} f = 0 \tag{27.135}$$

pour tout triangle (plein) T contenu dans Ω, alors f est holomorphe sur Ω.

Démonstration. Il est suffisant de prouver que f est holomorphe sur toute boule ouverte $B(a, r)$ inclue dans Ω. Nous posons, pour tout $z \in B(a, r)$,

$$F(z) = \int_{[p,z]} f, \tag{27.136}$$

et nous considérons le chemin triangulaire $a \to z \to z + h \to a$ 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

$$0 = \int_{a \to z} f + \int_{z \to z+h} f + \int_{z+h \to a} f, \tag{27.137}$$

c'est-à-dire

$$F(z + h) - F(z) = \int_{z \to z+h} f. \tag{27.138}$$

13. Définition 15.227.

En paramétrant le chemin par $z+th$ avec $t\in[0,1]$, et en tenant compte de la remarque 21.44,

$$F'(z)=\lim_{h\to 0}\frac{F(z+h)-F(z)}{h} \tag{27.139a}$$

$$=\lim_{h\to 0}\frac{1}{h}\int_0^1 f(z+th)hdt, \tag{27.139b}$$

ce qui prouve que F est dérivable et $F'=f$. Par définition (13.244), F est holomorphe, et donc C^∞ par le théorème 27.15. Du coup f est également C^∞ et donc en particulier holomorphe. □

27.2.1 Mesure de Radon

Définition 27.32.
*Une **mesure de Radon** sur un compact K de $\mathbb{C}$ est une forme linéaire continue sur $C(K)$. Si μ est une mesure de Radon, on définit la **transformée de Cauchy** de μ par*

$$\begin{aligned}\hat{\mu}\colon \mathbb{C}\backslash K&\to\mathbb{C}\\ z&\mapsto -\frac{1}{\pi}\mu\left(\frac{1}{\xi-z}\right).\end{aligned} \tag{27.140}$$

Théorème 27.33.
Si μ est une mesure de Radon sur K alors $\hat{\mu}$ est infiniment $\mathbb{C}$-dérivable sur $\Omega=\mathbb{C}\backslash K$ et nous avons

$$\hat{\mu}^{(n)}(z)=-\frac{n!}{\pi}\mu\left(\frac{1}{(\xi-z)^{n+1}}\right). \tag{27.141}$$

Cette théorie permet de fournir une démonstration plus technologique du corollaire 27.29.

Lemme 27.34.
Si f est holomorphe sur Ω et si B est une boule fermée dans Ω alors pour tout $z\in\mathrm{Int}(B)$ nous avons

$$f^{(k)}(z)=\frac{k!}{2i\pi}\int_{\partial B}\frac{f(\xi)}{(\xi-z)^{k+1}}d\xi. \tag{27.142}$$

Démonstration. Appliquer le théorème 27.33 à la mesure de Radon

$$\mu(\phi)=\int_{\partial B}\phi(\xi)d\xi. \tag{27.143}$$

□

Tout ce petit monde à propos de la mesure de Radon permet également de redémontrer que

$$\left(\frac{1}{2\pi i}\int_{\partial B}\frac{f(\xi)}{\xi^{n+1}}\right)=f^{(n)}(0)/n!, \tag{27.144}$$

comme nous l'avons déjà fait autour de l'équation (27.134). Nous utilisons le théorème de Radon 27.33 à la mesure

$$\mu(\phi)=\int_{\partial B}\phi(\xi)d\xi. \tag{27.145}$$

La transformée de Cauchy est

$$\hat{\mu}(z)=-\frac{1}{\pi}\mu\left(\frac{1}{\xi-z}\right)=-\frac{1}{\pi}\int_{\partial B}\frac{1}{\xi-z}d\xi, \tag{27.146}$$

et le théorème assure que

$$\hat{\mu}^{(n)}(z)=-\frac{n!}{\pi}\mu\left(\frac{1}{(\xi-z)^{n+1}}\right)=-\frac{n!}{\pi}\int_{\partial B}\frac{1}{(\xi-z)^{n+1}}d\xi. \tag{27.147}$$

En comparant la formule (27.146) avec la formule de Cauchy nous voyons que $\hat{\mu}(z) = -2if(z)$. Par conséquent

$$f^{(n)}(z) = -\frac{1}{2i}\hat{\mu}^{(n)}(z) = \frac{n!}{2\pi i}\int_{\partial B}\frac{1}{(\xi - z)^{n+1}}d\xi, \tag{27.148}$$

et

$$f^{(n)}(0) = \frac{n!}{2\pi i}\int_{\partial B}\frac{1}{\xi^{n+1}}d\xi. \tag{27.149}$$

27.3 Conditions équivalentes à l'holomorphie

Nous nous proposons de lister les conditions que nous avons vues être équivalentes à l'holomorphie.

Théorème 27.35.
Soient Ω un ouvert de $\mathbb{C}$ et $f\colon \Omega \to \mathbb{C}$ une fonction continue. Les conditions suivantes sont équivalentes.

(1) f est holomorphe.
(2) Pour tout triangle (plein) T contenu dans Ω, $\int_T f = 0$.
(3) f est $\mathbb{C}$-dérivable.
(4) f est C^∞
(5) $\frac{\partial f}{\partial \bar{z}} = 0$; 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 Ω nous avons

$$f(a) = \frac{1}{2\pi i}\int_{\partial B(a,r)}\frac{f(z)}{z-a}dz. \tag{27.150}$$

La fonction f est holomorphe 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,

$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n. \tag{27.151}$$

Dans ce cas, f est holomorphe sur toute la boule.

Démonstration. (1) implique (2) est le lemme de Goursat 27.5. (2) implique (1) est le théorème de Morera 27.31.

(3) est la définition de l'holomorphie, définition 13.244.

(4) implique (1) est un a fortiori sur la définition. (1) implique (4) est contenu dans le théorème de développement en série entière 27.15.

L'équivalence entre (5) et l'holomorphie est le théorème 27.2.

L'équivalence entre (6) et (1) est la proposition 27.18.

L'équivalence entre (1) et (7) est d'une part le théorème 27.15 et d'autre part le corollaire 27.16.

En ce qui concerne la dernière affirmation, si f est holomorphe en z_0, alors le théorème 27.15(1) donne la série. Si au contraire nous avons la série, la proposition 27.7 nous donne le résultat. □

27.4 Singularités, pôles et méromorphe

Définition 27.36.
*Si f est holomorphe sur un ouvert Ω, alors une **singularité** de f est un point isolé du bord de Ω.*

*(1) La singularité est **effaçable** si la fonction f s'y prolonge en une fonction holomorphe.*
*(2) La singularité Z est un **pôle** d'ordre k de f si elle n'est pas effaçable et si la fonction $z \mapsto (z - Z)^k f(z)$ se prolonge en une fonction holomorphe en Z.*

Exemple 27.37
La fonction

$$z \mapsto \frac{\sin(z)}{z} \tag{27.152}$$

n'est pas définie en $z = 0$, mais elle s'y prolonge en une fonction continue en posant $f(0) = 1$. △

Proposition 27.38.
Une singularité de f est un pôle si et seulement si

$$\lim_{z \to Z} f(z) = \infty. \tag{27.153}$$

Le théorème suivant compète la proposition 27.25.

Théorème 27.39 (Prolongement de Riemann)**.**
Soient $f : \Omega \to \mathbb{C}$ et Z une singularité de f. Nous avons équivalence de

(1) la singularité Z est effaçable ;
(2) f possède un prolongement continue en Z ;
(3) il existe un voisinage épointé de Z sur lequel f est bornée ;
(4) $\lim_{z \to Z}(z - Z)f(z) = 0$.

Définition 27.40 (Fonction méromorphe[400])**.**
*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 $\mathcal{U} \backslash \{p_i\}$ et si chaque p_i est un point régulier ou un pôle de f, alors nous disons que f est **méromorphe** sur $\mathcal{U}$.*

Proposition 27.41.
Soient Ω un ouvert de $\mathbb{C}$ et $f_n : \Omega \to \mathbb{C}$ une suite de fonctions telles que pour tout compact K de Ω il existe $N_K \geqslant 0$ tel que

(1) f_n n'a pas de pôle dans K dès que $n \geqslant N_K$;
(2) la série $\sum_{n \geqslant N_K} f_n$ converge uniformément sur K.

Alors

(1) La fonction

$$f(z) = \sum_{n=0}^{\infty} f_n(z) \tag{27.154}$$

est méromorphe sur Ω et ses pôles sont l'union de ceux des f_n.

(2) Nous pouvons permuter la somme et la dérivée :

$$f'(z) = \sum_{n=0}^{\infty} f_n'(z). \tag{27.155}$$

Théorème 27.42 (Série de Laurent)**.**
*Soient C une couronne de rayons $r_1 < r_2$ centrée en zéro et une fonction f holomorphe dans cette couronne. Alors nous avons la **série de Laurent***

$$f(z) = \sum_{n \in \mathbb{Z}} a_n z^n. \tag{27.156}$$

(1) Cette série converge uniformément sur tout compact de C.
(2) Les coefficients sont donnés par

$$a_n = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z^{n+1}} dz \tag{27.157}$$

où γ est un cercle centré en zéro.

(3) Ce développement en série est unique.
(4) La valeur des a_n ne dépend pas du choix du rayon du cercle γ.

27.5 Fonctions d'Euler

Théorème 27.43 (Prolongement méromorphe de la fonction Γ d'Euler[43]).
Nous considérons la formule

$$\Gamma(z) = \int_0^\infty e^{-t}t^{z-1}dt. \tag{27.158}$$

Alors

(1) Cette formule définit une fonction holomorphe sur

$$\mathcal{P} = \{z \in \mathbb{C} \text{ tel que } \Re(z) > 0\}. \tag{27.159}$$

(2) La fonction $\Gamma\colon \mathcal{P} \to \mathbb{C}$ admet un unique prolongement méromorphe sur $\mathbb{C}$, lequel a des pôles sur les entiers négatifs.

Démonstration. **Holomorphie sous l'intégrale** Pour étudier l'holomorphie de la fonction Γ sur $\mathcal{P}$ nous utilisons le théorème 27.28.

Nous considérons la fonction

$$\begin{aligned} g\colon \mathcal{P} \times \mathbb{R}^+ &\to \mathbb{C} \\ (z,t) &\mapsto e^{-t}z^{z-1} \end{aligned} \tag{27.160}$$

et nous commençons par montrer que c'est holomorphe en z pour chaque $t > 0$ fixé. Nous le vérifions par le critère de $\partial_{\bar{z}f=0}$ [14] et en nous souvenant que $t^i = e^{\ln(t^i)} = e^{i\ln(t)}$. Nous obtenons rapidement que

$$\frac{\partial g}{\partial \bar{z}} = 0. \tag{27.161}$$

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 $|f(z,t) \leqslant g(t)|$. Pour cela nous majorons séparément les parties $t \in]0,1[$ et $t \geqslant 1$.

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

$$e^{-t}t^{z-1} = e^{-t}e^{(z-1)\ln(t)}, \tag{27.162}$$

de telle façon à que que

$$\begin{aligned} |e^{-t}t^{z-1}| &\leqslant |e^{(x-1+iy)\ln(t)}| && \text{(27.163a)} \\ &= |e^{(\Re(z)-1)\ln(t)}| && \text{(27.163b)} \\ &= |t^{\Re(z)-1}| && \text{(27.163c)} \\ &\leqslant |t^{\epsilon-1}| && \text{(27.163d)} \\ &= \frac{1}{t^{1-\epsilon}}. && \text{(27.163e)} \end{aligned}$$

Cette dernière fonction est intégrable sur $]0,1[$.

Nous considérons maintenant $t \geqslant 1$. Dans ce cas nous avons

$$|e^{-t}z^{z-1}| = e^{-t}t^{\Re(z)-1} \leqslant e^{-t}t^{M-1}. \tag{27.164}$$

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

$$g_K(t) = \begin{cases} \frac{1}{t^{1-\epsilon}} & \text{si } t < 1 \\ \text{borné} & \text{si } 1 \leqslant t \leqslant b \\ e^{-t}t^{M-1} & \text{si } t > b. \end{cases} \tag{27.165}$$

14. Théorème 27.2.

Cela est une fonction intégrable sur $]0\infty,[$ et qui majore f uniformément en z sur le compact K de $\mathcal{P}$. Le théorème 27.28 nous permet donc de conclure que

$$\Gamma(z) = \int_0^\infty f(z,t)dt \tag{27.166}$$

est holomorphe en z sur $\mathcal{P}$ et que

$$\Gamma'(z) = \int_0^\infty \frac{\partial f}{\partial z}(z,t)dt. \tag{27.167}$$

En deux morceaux Nous passons maintenant à la seconde partie du théorème. Pour $z \in \mathcal{P}$ nous coupons l'intégrale en deux :

$$\Gamma(z) = \int_0^1 e^{-t}t^{z-1}dt + \int_1^\infty e^{-t}t^{z-1}dt \tag{27.168}$$

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

$$\int_0^1 \sum_{n=0}^\infty \frac{(-1)^n}{n!}t^{n+z-1}dt. \tag{27.169}$$

Nous allons permuter la somme avec l'intégrale à l'aide du théorème de Fubini 15.259 en posant la fonction

$$g(n,t) = \frac{(-1)^n}{n!}t^{n+z-1} \tag{27.170}$$

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

$$\int_0^1 \int_{\mathbb{N}} g(n,t)dndt. \tag{27.171}$$

Pour permuter il suffit de prouver que $|g|$ est intégrable pour la mesure produit, c'est-à-dire que

$$\int_0^1 \int_{\mathbb{N}} \left|\frac{(-1)^n}{n!}t^{n+z-1}\right| < \infty. \tag{27.172}$$

Nous avons $|t^z = t^{\Re(z)}|$, donc

$$\sum_{n=0}^\infty \left|\frac{t^{n+z-1}}{n!}\right| = t^{\Re(z)-1}\sum_{n=0}^\infty \frac{t^n}{n!} = t^{\Re(z)-1}e^t. \tag{27.173}$$

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. 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 :

$$\int_0^1 e^{-t}t^{z-1}dt = \sum_{n=0}^\infty \int_0^1 \frac{(-1)^n}{n!}t^{n+z-1}dt = \sum_{n=0}^\infty \frac{(-1)^n}{n!}[t^{n+z}]_0^1 = \sum_{n=0}^\infty \frac{(-1)^n}{n!(n+z)}. \tag{27.174}$$

Nous avons donc l'intéressante formule suivante, valable pour tout $z \in \mathcal{P}$:

$$\Gamma(z) = \sum_{n=0}^\infty \frac{(-1)^n}{n!(n-z)} + \int_1^\infty e^{-t}t^{z-1}dt. \tag{27.175}$$

Prolongation de la première partie Nous voudrions montrer maintenant que la fonction

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{n!(n-z)} \tag{27.176}$$

est méromorphe sur $\mathbb{C}$ avec des pôles en les entiers négatifs. Pour cela nous considérons la suite de fonctions

$$f_n(z)=\frac{(-1)^n}{n!(z+n)} \tag{27.177}$$

et nous allons utiliser la proposition 27.41. Si $n\geqslant 0$, 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 $n\geqslant N_K+1$, la fonction f_n n'a pas de pôle dans K et de plus pour tout $z\in K$ nous avons

$$|z+n|=|z-(-z)|\geqslant\big|n-|z|\big|\geqslant n-|z|\geqslant n-N_K, \tag{27.178}$$

et par conséquent

$$|f_n(z)|\leqslant\frac{1}{n!(n-N)}, \tag{27.179}$$

ou pour le dire de façon plus snob :

$$\|f_n\|_{\infty,K}\leqslant\frac{1}{n!(n-N)}, \tag{27.180}$$

dont la série converge. Cela signifie que la série $\sum_{n\geqslant N}f_n$ converge normalement [15] sur K, donc la fonction

$$f(z)=\sum_{n=0}^{\infty}f_n(z) \tag{27.181}$$

est une fonction méromorphe dont les pôles sont ceux des f_n, c'est-à-dire les entiers négatifs (proposition 27.41).

La seconde partie Nous allons à présent prouver que la fonction

$$g(z)=\int_1^{\infty}e^{-t}t^{z-1}dt \tag{27.182}$$

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'holomorphie sous l'intégrale 27.28. D'abord pour z_0 fixé dans $\mathbb{C}$ nous avons

$$\int_1^{\infty}|e^{-t}t^{z_0-1}|\leqslant\int_1^{\infty}e^{-t}t^{\Re(z_0)-1}dt, \tag{27.183}$$

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 :

$$\frac{1}{2}\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)t_0^{x+iy-1}=0. \tag{27.184}$$

Et enfin si K est compact dans $\mathbb{C}$ nous avons

$$|f(z,t)|=|e^{-t}t^{z-1}|=e^{-t}|t^{\Re(z)-1}|\leqslant e^{-t}t^{M-1} \tag{27.185}$$

où $M=\max_{z\in K}\Re(z)$. Nous en déduisons que la fonction

$$z\mapsto\int_1^{\infty}e^{-t}t^{z-1}dt \tag{27.186}$$

est une fonction holomorphe sur $\mathbb{C}$.

Conclusion Au final nous avons prouvé que la fonction Γ d'Euler admet le prolongement méromorphe sur $\mathbb{C}$ donné par

$$\Gamma(z)=\sum_{n=0}^{\infty}\frac{(-1)^n}{n!(z+n)}+\int_1^{\infty}e^{-t}t^{z-1}dt. \tag{27.187}$$

□

15. Définition 12.60.

27.5.1 Euler et factorielle

Proposition 27.44.
Nous avons la formule $\Gamma(n) = (n-1)!$ *pour tout* $n \in \mathbb{N}$.

Démonstration. Nous partons de la formule

$$\Gamma(n) = \int_0^\infty e^{-t} t^{n-1} dt \tag{27.188}$$

que nous intégrons par partie en posant

$$\begin{aligned} u &= t^{n-1} & u' &= (n-1)t^{n-1} \\ v &= e^{-t} & v' &= -e^{-t}. \end{aligned} \tag{27.189}$$

Les termes au bord s'annulent (ici il y a un passage à la limite qui n'est pas écrit) et nous trouvons

$$\Gamma(n) = \int_0^\infty (n-1)e^{-t}t^{n-2}dt = (n-1)\Gamma(n-1). \tag{27.190}$$

Pour conclure il suffit de remarquer que

$$\Gamma(1) = \int_0^\infty = -[e^{-t}]_0^\infty = 1. \tag{27.191}$$

□

27.6 Partition d'un entier en parts fixées

Proposition 27.45 ([43])**.**
Soient $a_1, \dots, a_k \in \mathbb{N}^*$ *des entiers premiers entre eux dans leur ensemble. Pour* $n \geqslant 1$ *nous posons*

$$u_n = \text{Card}\left\{ (x_1, \dots, x_k) \in \mathbb{N}^* \text{ tel que } \sum_{i=1}^k a_i x_i = n \right\}, \tag{27.192}$$

et $u_0 = 1$.

Alors nous avons l'équivalence de suite (pour $n \to \infty$*) :*

$$u_n \sim \frac{1}{a_1 \dots a_k} \frac{n^{k-1}}{(k-1)!}. \tag{27.193}$$

Démonstration. Pour chacun des $i \in \{1, \dots, k\}$ nous considérons la série entière

$$\sum_{x=0}^\infty z^{xa_i} = \sum_k (z^{a_i})^x. \tag{27.194}$$

Étant donné que $|z^{a_i}| < 1$ si et seulement si $|z < 1|$, cette série a un rayon de convergence égal à 1. Nous allons calculer le produit de Cauchy de ces k séries, en nous souvenant que le théorème 16.15 nous assure que la série résultante aura un rayon de convergence au moins égal à 1 et vaudra le produit des différentes séries.

Le coefficient de z^n dans cette série vaut

$$\sum_{\substack{x \in \mathbb{N}^k \\ \sum x_i a_i = n}} 1 = u_n \tag{27.195}$$

parce que dans chacune des séries, le coefficient de tous les z^{xa_i} est 1. Nous définissons la fonction

$$f(z) = \sum_{n=0}^\infty u_n z^n = \prod_{i=1}^k \left(\sum_{x=0}^\infty z^{xa_i} \right) = \prod_{i=1}^k \frac{1}{1 - z^{a_i}}. \tag{27.196}$$

La fonction f existe sur $|z|<1$ parce que nous venons de voir qu'elle peut s'exprimer comme un produit de Cauchy ; et la dernière égalité est simplement la somme de la série harmonique. D'autre part la fonction f est la série génératrice de la suite (u_n).

Nous sommes en présence d'une fonction ayant des pôles aux racines $a_1,\dots,a_k$[e] de l'unité. Étant donné que 1 est une racine de l'unité de tous les ordres, le pôle en $z=1$ est de multiplicité k. Les autres pôles sont de multiplicité strictement inférieure ; en effet soit $\omega\in\mathbb{C}$ tel que $\omega^{a_i}=1$ pour tout i. Alors Bezout [16] nous donne des entiers $v_i\in\mathbb{Z}$ tels que $\sum_i v_i a_i=1$. Alors nous avons

$$\omega=\omega^{\sum_{v_i a_i}}=\prod_{i=1}^{k}(\omega^{a_i})^{v_i}=1. \tag{27.197}$$

Donc nous voyons que 1 est le seul à être racine de tous les ordres en même temps. Nous notons

$$P=\{\omega_1,\dots,\omega_p\} \tag{27.198}$$

l'ensemble des pôles avec $\omega_1=1$. Par ailleurs la fonction f est une fraction rationnelle dont nous connaissons les racines du dénominateur (ce sont les ω_i) et à peu près leurs ordres. Nous utilisons le truc de la décomposition en fractions simples en séparant le terme de puissance k qui n'existe que pour la racine $\omega_1=1$:

$$f(z)=\frac{\alpha}{(1-z)^k}+\sum_{i=1}^{p}\sum_{j=1}^{k-1}\frac{c_{ij}}{(\omega_i-z)^j}. \tag{27.199}$$

Ce développement est valable pour tout $|z|<1$. Nous considérons maintenant $\omega\in P$ et $j\in\mathbb{N}$ et nous étudions la fonction

$$g(z)=\frac{1}{\omega-z}. \tag{27.200}$$

Un rapide calcul (par exemple par récurrence) montre que

$$g^{(k)}(z)=\frac{k!}{(\omega-z)^{k+1}}, \tag{27.201}$$

et étant donné que $|\omega|=1$ nous pouvons écrire la série

$$\frac{1}{\omega-z}=\sum_{k=0}^{\infty}\frac{z^k}{\omega^{k+1}}, \tag{27.202}$$

valable pour $|z|<1$. Ce qui nous intéresse, c'est d'exprimer une série pour $1/(\omega-z)^j$; et voyant (27.201), nous voyons qu'il suffit de calculer les dérivées de la série de g. Nous dériver terme à terme à l'intérieur du rayon de convergence. Avec quelques abus d'écriture, et en utilisant la bête

16. Théorème 3.13.

formule (16.79) nous avons [17]

$$\frac{1}{(\omega-z)^j} = \frac{g^{(j-1)}(z)}{(j-1)!} \tag{27.203a}$$

$$= \frac{1}{(j-1)!}\left(\frac{1}{\omega-z}\right)^{(j-1)} \tag{27.203b}$$

$$= \frac{1}{(j-1)!}\sum_{k=0}^{\infty}\frac{1}{\omega^{k+1}}(z^k)^{(j-1)} \tag{27.203c}$$

$$= \sum_{k=j-\infty}\frac{1}{(j-1)!}\frac{1}{\omega^{k+1}}\frac{k!}{(k-j+1)!}z^{k-j+1} \tag{27.203d}$$

$$= \sum_{n=0}^{\infty}\frac{1}{\omega^{n+j}}\frac{(n+j-1)!}{n!(j-1)!}z^n \tag{27.203e}$$

$$= \sum_{n=0}^{\infty}\frac{1}{\omega^{n+j}}\binom{n+j-1}{n}z^n. \tag{27.203f}$$

Nous pouvons utiliser cela pour récrire la formule (27.199) de façon considérablement plus compliquée :

$$f(z) = \alpha\sum_{n=0}^{\infty}\binom{n+j-1}{n}z^n + \sum_{i=1}^{p}\sum_{j=1}^{k-1}\sum_{n=0}^{\infty}c_{ij}\binom{n+j-1}{n}\frac{z^n}{\omega_i^{n+j}}. \tag{27.204}$$

Mais nous savons que ce f est la série génératrice de la suite (u_n) et que nous pouvons donc utiliser la formule (16.347) pour exprimer les nombres u_l : u_l est simplement le coefficient de z^l divisé par $l!$. C'est-à-dire

$$u_l = \alpha\binom{l+k-1}{l} + \sum_{i=1}^{p}\sum_{j=1}^{k-1}c_{ij}\binom{l+j-1}{l}\frac{1}{\omega_i^{l+j}}. \tag{27.205}$$

Notre boulot est d'examiner le comportement de cela lorsque $l\to\infty$, c'est-à-dire regarder quels sont les puissances de l en présence. Notons que

En ce qui concerne le premier terme, la puissance dominante dans le coefficient binomial est l^{k-1}. Dans les autres termes [18], c'est l^{j-1} qui est de degré moins grand. Donc le comportement de u_l en terme de l est

$$u_l \sim \alpha\frac{l^{k-1}}{(k-1)!}. \tag{27.206}$$

Il nous reste à voir ce que vaut α. Pour cela nous repartons de l'expression 27.196 que nous écrivons sous la forme

$$(1-z)^k f(z) = \prod_{i=1}^{k}\frac{1-z}{1-z^{a_i}}. \tag{27.207}$$

Nous reconnaissons l'inverse d'une somme harmonique partielle :

$$(1-z)^k f(z) = \prod_{i=1}^{k}\frac{1}{1+z+z^2+\cdots+z^{a_i-1}}. \tag{27.208}$$

Par ailleurs, nous ne savons pas si $f(1)$ existe parce que son rayon de convergence n'est que de 1 ; et nous savons même qu'elle n'existe pas (parce que ce serait la somme des u_n). Mais nous

17. À ce niveau j'ai pas exactement le même coefficient binomial que dans [43], mais je n'exclus absolument pas que ce soit moi qui me trompe. Écrivez-moi si vous pouvez infirmer ou confirmer l'erreur. Quoi qu'il en soit, cela ne change pas le résultat asymptotique que nous cherchons.

18. Attention : les termes $i=1$ ont $\omega_1=1$ et il n'est donc pas possible de conclure simplement en disant que $\omega_i^{l-j}\to 0$ pour $l\to\infty$; bien que cela soit vrai pour tous les $i\neq 1$.

savons aussi que le pôle de plus grande multiplicité de f est en $z=1$ et est de multiplicité k. Donc $(1-z)^k f(z)$ devrait converger pour $z\to 1$. Pour tout $|z<1|$ nous avons

$$(1-z)^k f(z)=\alpha+\sum_{i=1}^{p}\sum_{j=1}^{k-1}c_{ij}\frac{(1-z)^k}{(\omega_i-z)^j}. \tag{27.209}$$

Lorsque $z\to 1$, tous les termes des sommes tendent vers zéro, y compris ceux avec $i=1$ parce que $j<k$. Il reste donc

$$\lim_{z\to 0}(1-z)^k f(z)=\alpha. \tag{27.210}$$

En calculant la même limite avec (27.208) nous trouvons

$$\lim_{z\to 1}(1-z)^k f(z)=\lim_{z\to 1}\prod_{i=1}^{k}\frac{1}{1+z+z^2+\cdots+z^{a_i-1}}=\frac{1}{a_1\dots a_k}. \tag{27.211}$$

Donc

$$\alpha=\frac{1}{a_1\dots a_k}, \tag{27.212}$$

et le résultat est prouvé.

□

27.7 Exponentielle et logarithme complexe

27.7.1 Propriétés de l'exponentielle

Proposition 27.46.
Soit $z\in\mathbb{C}$ fixé. La fonction

$$\begin{aligned}E\colon\mathbb{R}&\to\mathbb{C}\\ t&\mapsto e^{tz}\end{aligned} \tag{27.213}$$

est C^∞, sa dérivée est

$$E'(t)=ze^{tz}. \tag{27.214}$$

La fonction E est développable en série entière (voir définition 16.98) sur $\mathbb{R}$ en $t=0$ et

$$e^{tz}=\sum_{n=0}^{\infty}\frac{z^n}{n!}t^n. \tag{27.215}$$

Démonstration. Nous fixons $z\in\mathbb{C}$. Par définition 19.20, la série suivante est e^{tz} :

$$f(t)=\sum_{n=0}^{\infty}\frac{z^n}{n!}t^n. \tag{27.216}$$

Cette série a un rayon de convergence infini et la fonction f est donc C^∞ sur $\mathbb{R}$. Nous pouvons la dériver terme à terme :

$$f'(t)=\sum_{n=1}^{\infty}\frac{z^n}{n!}nt^{n-1}=z\sum_{n=1}^{\infty}\frac{z^{n-1}}{(n-1)!}t^{n-1}=ze^{tz}. \tag{27.217}$$

□

Théorème 27.47.
La fonction exponentielle vérifie les propriétés suivantes.

(1) exp *est holomorphe*[19].

(2) $(e^z)'=e^z$.

19. Définition 13.244.

(3) L'exponentielle est développable en série entière,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \tag{27.218}$$

et la série converge normalement sur tout compact de $\mathbb{C}$.

Démonstration. En tant que application $E\colon \mathbb{R}^2 \to \mathbb{C}$, la fonction

$$E(x,y) = e^x(\cos y + i \sin y) \tag{27.219}$$

est C^∞. De plus nous avons

$$\frac{\partial E}{\partial x}(x,y) = e^{x+iy} = E(x,y) \tag{27.220a}$$

$$\frac{\partial E}{\partial y}(x,y) = iE(x,y), \tag{27.220b}$$

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 $r/n!$ nous savons que la série $\sum z^n/n!$ converge normalement sur le compact $B(0,r)$. □

27.7.2 Intégrale de Fresnel

Nous allons calculer l'**intégrale de Fresnel**

$$\int_0^\infty e^{-ix^2} dx = \frac{\sqrt{\pi}}{2} e^{-i\pi/4} \tag{27.221}$$

en suivant la démarche présentée par Wikipédia[401]. Nous commençons par prouver que l'intégrale est convergente en nous contentant de justifier la convergence de

$$\int_0^\infty \sin(x^2) dx. \tag{27.222}$$

Pour chaque $a > 0$ fixé, l'intégrale $\int_0^a \sin(x^2)dx$ ne pose pas de problèmes. Le lemme 21.197 nous permet de passer à la limite ; nous devons donc seulement calculer

$$\lim_{b\to\infty} \int_a^b \sin(x^2) dx \tag{27.223}$$

où a est une constante strictement positive. Nous effectuons une intégration par partie en posant

$$u = \frac{1}{x} \qquad u' = -\frac{1}{x^2} \tag{27.224a}$$

$$v' = x\sin(x) \qquad v = \frac{1-\cos(x)}{2}. \tag{27.224b}$$

Notons que la primitive v a été choisie pour avoir $v(0) = 0$. Nous avons

$$\int_a^b \sin(x^2) dx = \left[\frac{1-\cos(x^2)}{2x}\right]_a^b - \int_a^b \frac{\cos(x^2)-1}{2x^2} dx \tag{27.225}$$

Pour le premier terme nous avons

$$\lim_{b\to\infty} \left[\frac{1-\cos(x^2)}{2x}\right]_a^b = \lim_{b\to\infty} \frac{1-\cos(b^2)}{2b} - \frac{1-\cos(a^2)}{2a} = -\frac{1-\cos(a^2)}{2a}. \tag{27.226}$$

C'est borné. Pour le second terme de (27.225), la fonction

$$\frac{\cos(x^2)-1}{2x^2} \tag{27.227}$$

est majorée par la fonction $1/x^2$ qui est intégrable entre a et ∞.

Nous allons calculer l'intégrale demandée en passant par la fonction

$$f(x) = e^{-z^2} \tag{27.228}$$

définie sur le plan complexe. Nous l'intégrons sur le chemin $\gamma = \gamma_1 + \gamma_2 - \gamma_3$ indiqué à la figure 27.1.

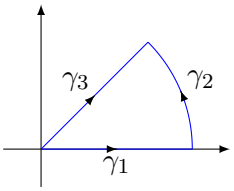


FIGURE 27.1 – Chemin d'intégration pour l'intégrale de Fresnel

Ces chemins sont donnés par

$$\begin{aligned}\gamma_1 : [0, R] &\to \mathbb{C} \\ t &\mapsto t,\end{aligned} \tag{27.229}$$

$$\begin{aligned}\gamma_2 : [0, \frac{\pi}{4}] &\to \mathbb{C} \\ t &\mapsto Re^{it},\end{aligned} \tag{27.230}$$

$$\begin{aligned}\gamma_3 : [0, R] &\to \mathbb{C} \\ t &\mapsto te^{i\pi/4}.\end{aligned} \tag{27.231}$$

Tout d'abord la fonction f est bien holomorphe par le critère du théorème 27.2. 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

$$0 = \int_\gamma f = \underbrace{\int_0^R e^{-t^2} dt}_{I_1(R)} + \underbrace{\int_0^{\pi/4} e^{-R^2 e^{2it}} Rie^{it} dt}_{I_2(R)} + \underbrace{\int_0^R e^{-t^2 e^{i\pi/2}} e^{i\pi/4} dt}_{I_3(R)}. \tag{27.232}$$

L'intégrale est nulle pour tout R en vertu de la proposition 27.6. L'intégrale I_1 est une gaussienne et nous avons

$$\lim_{R\to\infty} I_1(R) = \frac{\sqrt{\pi}}{2} \tag{27.233}$$

par l'exemple 15.262. Nous montrons maintenant que $\lim_{R\to\infty} |I_2(R)| = 0$ [20]. D'abord nous majorons en prenant la norme puis nous effectuons le changement de variables $u = 2t$:

$$|I_2(R)| \leqslant \int_0^{\pi/4} Re^{-R^2\cos(2t)} dt \tag{27.234a}$$

$$= \frac{R}{2} \int_0^{\pi/2} e^{-R^2\cos(u)} du. \tag{27.234b}$$

20. Il y a moyen de démontrer cela via le lemme de Jordan[402]. Nous donnons ici une démonstration moins technologique.

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 $\cos(u)\geqslant 1-\frac{2}{\pi}u$ et par conséquent

$$e^{-R^2\cos(u)}\leqslant e^{-R^2(1-\frac{2}{\pi}u)}=e^{R^2(\frac{2}{\pi}u-1)}. \tag{27.235}$$

Nous effectuons l'intégrale

$$\begin{aligned}|I_2(R)|&\leqslant\frac{R}{2}\int_0^{\pi/2}e^{-R^2}e^{\frac{2R^2}{\pi}u}du &(27.236a)\\ &=\frac{R}{2}e^{-R^2}\left[\frac{\pi}{2R^2}e^{2R^2u/\pi}\right]_0^{\pi/2} &(27.236b)\\ &=\frac{\pi}{4R}-\frac{\pi e^{-R^2}}{4R}, &(27.236c)\end{aligned}$$

et nous avons bien $\lim_{R\to\infty}|I_2(R)|=0$. Nous passons à la troisième intégrale. En tenant compte que $e^{i\pi/2}=i$, nous avons

$$\begin{aligned}I_3(R)&=-\int_0^R e^{-\gamma_3(t)^2}e^{i\pi/4}dt &(27.237a)\\ &=-\frac{1+i}{\sqrt{2}}\int_0^R e^{-t^2}e^{2i\pi/4} &(27.237b)\\ &=-\frac{1+i}{\sqrt{2}}\int_0^R e^{-it^2}. &(27.237c)\end{aligned}$$

En passant à la limite $R\to 0$, de l'équation (27.232) il ne reste que

$$0=\frac{\sqrt{2}}{2}-\frac{1+i}{\sqrt{2}}\int_0^\infty e^{-it^2}dt, \tag{27.238}$$

ce qui signifie que

$$\int_0^\infty e^{-it^2}dt=\frac{\sqrt{2\pi}}{2(1+i)}=\frac{\sqrt{\pi}}{2}e^{-i\pi/4}. \tag{27.239}$$

27.7.3 Logarithme complexe

27.7.3.1 La fonction argument

Nous savons la définition 19.20 de l'exponentielle complexe.

Définition 27.48.
*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.

Lemme 27.49.
Si z_1 et z_2 sont des logarithmes de α 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

$$e^ae^{bi}=1, \tag{27.240}$$

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 α. Alors $e^{z_1}=e^{z_2}$, donc[21] $e^{z_1-z_2}=1$, ce qui signifie que z_1-z_2 est un logarithme de 1. Donc il existe un $k\in\mathbb{Z}$ tel que $z_1-z_2=2ik\pi$. □

21. C'est facile de dire « donc ». Il faut surtout citer la proposition 19.23(2).

Remarque 27.50.
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 27.51 ([403]).
*Si $z \in \mathbb{C}^*$ nous définissons la **valeur principale** de son argument le nombre $\theta \in]-\pi, \pi]$ tel que*

$$z = |z|e^{i\theta} \tag{27.241}$$

Nous le notons $\arg(z)$.

27.52.
Il ne faut pas se ruer sur $\arg(x+iy) = \arctan(y/x)$. Pour rappel, la fonction arctan a été définie dans le théorème 19.43, 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 :

$$\arg(x+iy) = \begin{cases} \arctan(y/x) & \text{si } x > 0 \\ \pi + \arctan(y/x) & \text{si } x < 0 \text{ et } y \geqslant 0 \\ -\pi + \arctan(y/x) & \text{si } x < 0 \text{ et } y < 0 \\ \frac{\pi}{2} & \text{si } x = 0 \text{ et } y > 0 \\ \frac{-\pi}{2} & \text{si } x = 0 \text{ et } y < 0. \end{cases} \tag{27.242}$$

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 π. 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 $\arg(x+iy) = -\pi + \arctan(y/x)$.

27.53 (Les dérivées partielles de la fonction argument)**.**
Vu que nous en aurons besoin plusieurs fois, nous calculons maintenant les dérivées partielles de la fonction

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\to \mathbb{R} \\ (x,y) &\mapsto \arg(x+iy). \end{aligned} \tag{27.243}$$

Nous commençons par la dérivée $\partial_x\varphi(x,y)$. Et il y a de nombreux cas à séparer.

<u>$x > 0$</u> Nous avons

$$\frac{\partial\varphi}{\partial x}(x,y) = \lim_{\epsilon\to 0} \frac{\arctan(y/(x+\epsilon)) - \arctan(y/x)}{\epsilon}, \tag{27.244}$$

qui n'est autre que la dérivée de la fonction $x \mapsto \arctan(y/x)$. Nous pouvons la calculer facilement avec le théorème 19.43(2) :

$$\frac{\partial\varphi}{\partial x}(x,y) = -\frac{y}{x^2+y^2}. \tag{27.245}$$

<u>$x < 0$</u> Nous avons

$$\frac{\partial\varphi}{\partial x}(x,y) = \lim_{\epsilon\to 0} \frac{\pm\pi + \arctan(y/(x+\epsilon)) - \big(\pm\pi + \arctan(y/x)\big)}{\epsilon} \tag{27.246}$$

où les signes $\pm$ dépendent du signe de y. De toutes façons, les termes en π se simplifient et le calcul est le même que celui du cas $x > 0$. Encore une fois nous avons

$$\frac{\partial\varphi}{\partial x}(x,y) = -\frac{y}{x^2+y^2}. \tag{27.247}$$

<u>$x = 0$</u> Nous devons calculer

$$\frac{\partial\varphi}{\partial x}(0,y) = \lim_{\epsilon\to 0} \frac{\arg(\epsilon+iy) - \arg(iy)}{\epsilon}. \tag{27.248}$$

Il y a quatre cas d'après les signes de ϵ (séparer limite à gauche et à droite) et y.

Si $\epsilon > 0$ et $y > 0$ alors nous avons à faire le calcul

$$\lim_{\epsilon\to 0^+} \frac{\arctan(y/\epsilon) - \pi/2}{\epsilon} \tag{27.249}$$

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 [22]. Au final,

$$\frac{\partial\varphi}{\partial x}(0,y) = -\frac{1}{y}. \tag{27.250}$$

Nous avons calculé jusqu'ici :

$$\frac{\partial\varphi}{\partial x}(x,y) = \frac{-y}{x^2+y^2} \tag{27.251}$$

pour tout $(x,y) \in \mathbb{R}^2\backslash\{(0,0)\}$. En particulier vous avez noté que cette dérivée partielle est continue sur $\mathbb{R}^2\backslash\{(0,0)\}$.

Nous calculons à présent la dérivée partielle par rapport à y :

$$\frac{\partial\varphi}{\partial y}(x,y) = \lim_{\epsilon\to 0} \frac{\arg(x+iy+i\epsilon) - \arg(x+iy)}{\epsilon}. \tag{27.252}$$

$\underline{x > 0}$ Nous avons à calculer

$$\lim_{\epsilon\to 0} \frac{\arctan\frac{y+\epsilon}{x} - \arctan\frac{y}{x}}{\epsilon}, \tag{27.253}$$

qui n'est autre que la dérivée de la fonction $t \mapsto \arctan\frac{t}{x}$ en $t = y$. Résultat :

$$\frac{\partial\varphi}{\partial y}(x,y) = \frac{x}{x^2+y^2}. \tag{27.254}$$

$\underline{x < 0 \text{ et } y \neq 0}$ Le calcul à faire est :

$$\lim_{\epsilon\to 0} \frac{\pm\pi + \arctan\frac{y+\epsilon}{x} - \left(\pm\pi + \arctan\frac{y}{x}\right)}{\epsilon} \tag{27.255}$$

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 ϵ (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

$$\frac{\partial\varphi}{\partial y}(x,y) = \frac{x}{x^2+y^2} \tag{27.256}$$

$\underline{x < 0 \text{ et } y = 0}$ Vu que $x < 0$ nous avons $\arg(x) = \pi$ et nous devons calculer

$$\lim_{\epsilon\to 0} \frac{\arg(x+i\epsilon) - \pi}{\epsilon}. \tag{27.257}$$

La limite $\epsilon \to 0^+$ est classique et donne $1/x$.

Mais la limite $\epsilon \to 0^-$ n'existe pas :

$$\lim_{\epsilon\to 0^-} \frac{-\pi + \arctan(\epsilon/x) - \pi}{\epsilon} \tag{27.258}$$

n'existe pas.

Donc

$$\frac{\partial\varphi}{\partial y}(x,0) \tag{27.259}$$

n'existe pas pour $x < 0$.

22. 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.

$x = 0$ **et** $y \neq 0$ Le calcul est immédiat

$$\lim_{\epsilon \to 0} \frac{\arg(iy + i\epsilon) - \arg(iy)}{\epsilon} = 0, \tag{27.260}$$

donc

$$\frac{\partial \varphi}{\partial y}(0, y) = 0. \tag{27.261}$$

En ce qui concerne la continuité, nous avons que $\partial_y \varphi$ est continue partout sauf sur la demi-droite $\{(x, 0) \text{ tel que } x \leqslant 0\}$ où elle n'existe pas.

27.7.3.2 Une définition possible du logarithme

Définition 27.54.
Nous définissons la fonction ***logarithme*** *par*

$$\begin{aligned} \ln \colon \mathbb{C}^* &\to \mathbb{C} \\ z &\mapsto \ln\big(|z|\big) + i \arg(z) \end{aligned} \tag{27.262}$$

où le ln *à droite est le logarithme usuel sur* $\mathbb{R}^+$.

Remarque 27.55.
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 27.56.
Le nombre $\ln(z)$ *est un logarithme de* z.

Démonstration. Nous avons

$$e^{\ln(z)} = e^{\ln|z|} e^{i \arg(z)} = |z| e^{i \arg(z)} = z. \tag{27.263}$$

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. □

Notons que si on avait pris d'autres conventions pour définir arg, nous aurions eu d'autres définitions possibles de ln.

Exemple 27.57
Nous avons

$$\ln(-1) = \ln(1) + i \arg(-1). \tag{27.264}$$

Mais $\ln(1) = 0$ et $\arg(-1) = \pi$ (et non $-\pi$), donc

$$\ln(-1) = i\pi. \tag{27.265}$$

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
```

```
sage: f.integrate(x)
x |--> log(x)
```

tex/sage/sageSnip010.sage

△

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)

Lemme 27.58.
La fonction ln *n'est pas continue sur* $]-\infty,0]$.

Démonstration. Attention à bien comprendre l'énoncé. La fonction

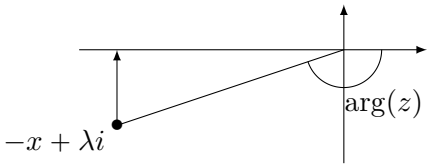
$$\begin{aligned} f\colon]-\infty,0[&\to \mathbb{C} \\ x &\mapsto \ln(x) \end{aligned} \tag{27.266}$$

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

$$\ln(-x) = \ln(x) + i\pi. \tag{27.267}$$

Cependant $\lim_{\substack{\lambda\to 0^-\\ \lambda\in\mathbb{R}}} \ln(-x+\lambda i)$ 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 π).



Donc

$$\lim_{\substack{\lambda\to 0^-\\ \lambda\in\mathbb{R}}} \ln(-x+\lambda i) = \lim \ln(|x+\lambda i|) + i\arg(-x+\lambda i) = \ln(|x|) - i\pi. \tag{27.268}$$

Nous n'avons donc pas continuité de la fonction logarithme comme fonction sur $\mathbb{C}^*$. □

Théorème 27.59.
La restriction

$$\ln\colon \mathbb{C}\backslash]-\infty,0] \to \mathbb{C} \tag{27.269}$$

est holomorphe.

Démonstration. Nous allons utiliser la proposition 13.249 et considérer la fonction

$$\begin{aligned} F\colon S &\to \mathbb{R}^2 \\ (x,y) &\mapsto \big(\ln(|x+iy|), \arg(x+iy)\big) \end{aligned} \tag{27.270}$$

où $S = \mathbb{R}^2\backslash\{(x,0) \text{ tel que } x \leqslant 0\}$. Nous devons vérifier que F est différentiable et que sa différentielle en un point de S est une similitude.

Nous posons

$$u(x,y) = \ln\big(\sqrt{x^2+y^2}\big) \tag{27.271}$$

et

$$v(x,y) = \arg(x+iy). \tag{27.272}$$

Les dérivées partielles de u ne sont pas très compliquées :

```
sage: var('x,y')
(x, y)
sage: u(x,y)=ln(sqrt(x**2+y**2))
sage: u.diff(x)
(x, y) |--> x/(x^2 + y^2)
```

tex/sage/sageSnip011.sage

c'est-à-dire

$$\frac{\partial u}{\partial x} = \frac{x}{x^2+y^2} \tag{27.273a}$$

$$\frac{\partial u}{\partial y} = \frac{y}{x^2+y^2}. \tag{27.273b}$$

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 27.52.

Nous avons déjà calculé les dérivées partielles de v dans 27.53, et nous avons vu qu'elles étaient continues sur $\mathbb{R}^2$ privé de la demi-droite.

Vu que les dérivées partielles sont continues, la proposition 13.223 nous dit que F est différentiable. La matrice de la différentielle est alors la matrice des dérivées partielles

$$\begin{pmatrix} \frac{x}{x^2+y^2} & \frac{y}{x^2+y^2} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}, \tag{27.274}$$

qui a la forme requise (13.612) pour que la proposition 13.249 nous assure que ln soit $\mathbb{C}$-dérivable, c'est-à-dire holomorphe. □

27.7.3.3 Pas plus de continuité

Bon. La fonction logarithme que nous avons définie est holomorphe sur $\mathbb{C}^*$ privé d'une demi-droite U. 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 27.60.
Il n'existe pas de fonctions continues $f\colon \mathbb{C}^* \to \mathbb{C}$ 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 27.49

$$f(z) = \ln(z) + 2ik(z)\pi \tag{27.275}$$

pour une certaine fonction $k\colon \mathbb{C}^* \to \mathbb{Z}$. 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

$$f(z) = \ln(z) + 2ik\pi \tag{27.276}$$

au moins pour tout $z \in \mathbb{C}^* \backslash U$. Une telle fonction ne peut pas être continue sur U parce que ln ne l'est pas. □

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 27.61.
Soit Ω 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 $f\colon \Omega \to \mathbb{C}$ telle que $e^{f(z)} = z$ pour tout $z \in \Omega$.

Démonstration. Encore une fois, pour tout $z \in \Omega$ nous avons

$$f(z) = \ln(z) + 2i\pi k(z) \tag{27.277}$$

pour une certaine fonction $k\colon \Omega \to \mathbb{Z}$. Sur $\Omega\backslash U$, la fonction ln est continue et k doit également l'être. Donc k est constante sur les composantes connexes de $\Omega\backslash U$.

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 Ω contient une couronne

$$\{z \in \mathbb{C} \text{ tel que } r - \delta \leqslant |z| \leqslant r + \delta\}. \tag{27.278}$$

Soit le point $x_0 = -r$. C'est un point de Ω contenu dans U. Nous allons prouver que $B(x_0,\delta)\backslash U$ est dans une seule composante connexe de Ω.

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 Ω et que $k(z_1) = k(z_2)$. Il existe donc $k \in \mathbb{Z}$ tel que

$$f(z) = \ln(z) + 2ik\pi \tag{27.279}$$

sur $B(x_0,\delta)\backslash U$. Une telle fonction f ne peut pas être continue. □

27.7.3.4 Pas d'unicité : autres déterminations de l'argument

27.62.
Nous avons pris la fonction d'argument $\arg\colon \mathbb{C} \to]-\pi,\pi]$. Il y en a évidemment beaucoup d'autres de possibles. Par exemple pour $\alpha \in \mathbb{R}$ nous pouvons considérer

$$\arg_{\alpha^+}\colon \mathbb{C} \to]\alpha, \alpha + 2\pi] \tag{27.280}$$

ou

$$\arg_{\alpha^-}\colon \mathbb{C} \to [\alpha, \alpha + 2\pi[. \tag{27.281}$$

En posant

$$\ln_{\alpha^\pm}(z) = \ln(|z|) + i \arg_{\alpha^\pm}(z) \tag{27.282}$$

nous avons une fonction réciproque de l'exponentielle définie sur $\mathbb{C}^*$ et holomorphe sur $\mathbb{C}^*$ privé d'une demi-droite D_α (dépendante de la valeur de α).

La différence entre $\ln_{\alpha^+}$ et $\ln_{\alpha^-}$ est seulement la valeur sur la demi-droite de non-holomorphie. L'une sera semi-continue d'un côté et l'autre, de l'autre côté.

Remarque 27.63.
La fonction $\arg_{0^-}$ a déjà été utilisée en 19.6.2 pour écrire un inverse de la fonction

$$\begin{aligned} \varphi\colon [0,2\pi[&\to S^1 \\ t &\mapsto e^{it}. \end{aligned} \tag{27.283}$$

Définition 27.64 ([404])**.**
Soit un ouvert $\Omega \subset \mathbb{C}^*$. Nous disons que la fonction $f\colon \Omega \to \mathbb{C}$ est une ***détermination*** *sur Ω si elle est continue et vérifie*

$$e^{f(z)} = z \tag{27.284}$$

pour tout $z \in \mathbb{C}$.

Les différents résultats vus jusqu'ici montrent qu'il n'existe pas de détermination du logarithme sur $\mathbb{C}^*$.

Définition 27.65.
*La **détermination principale** du logarithme est la restriction de notre logarithme 27.54*

$$\begin{aligned}\ln\colon \mathbb{C}^* &\to \mathbb{C}\\ z &\mapsto \ln(|z|)+i\arg(z)\end{aligned} \tag{27.285}$$

à l'ouvert $\mathbb{C}^*\backslash U$ où U est la partie $\Re(z)\leqslant 0$ de $\mathbb{C}$.

Remarque 27.66.
Beaucoup de sources[405] ne définissent pas $\ln_{\alpha^\pm}$ sur la droite D_α. C'est-à-dire qu'ils notent $\ln_\alpha$ notre fonction $\ln_{\alpha^+}$ restreinte à $\mathbb{C}^*\backslash D_\alpha$. Dans ce cas, les fonctions $\ln_{\alpha^+}$ et $\ln_{\alpha^-}$ sont identiques[23].

Cette remarque est importante parce que certains vont vous dire « le logarithme n'est pas définit 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 définit 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 « 2π et 0, c'est le même point ».

Nous préférons être bien clairs[24] sur ce point : notre fonction ln est parfaitement définie sur $\mathbb{C}^*$ et 2π 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 $\ln\colon \mathbb{R}^+\to\mathbb{R}$ que nous voulons prolonger sur $\mathbb{R}$, nous devons poser

$$\ln(-x)=\ln(-1)+\ln(x) \tag{27.286}$$

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 étaient possibles, comme $\ln(-1)=2017\pi$ par exemple.

27.7.3.5 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

$$\ln_{z_0}\colon B(z_0,|z_0|)\to\mathbb{C} \tag{27.287}$$

donnée par la série.

23. Cela n'est pas tout à fait évident ; vous devriez y penser.
24. Est-ce qu'il faut vraiment un pluriel ici ?

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.

27.7.3.6 Pas d'unicité : laquelle choisir ?

Bon. Pour chaque demi-droite D nous avons une détermination du logarithme sur $\mathbb{C}^*\backslash D$. 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.

Si vous avez surtout besoin d'holomorphie, et que vous en avez besoin sur un grand domaine, vous devriez choisir une détermination sur un des ensembles $\mathbb{C}^*\backslash D_\alpha$ en choisissant α 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 aurions besoin d'holomorphie, nous préciserons que nous considérons la restriction.

27.7.3.7 Logarithme comme primitive

Tout le monde sait que le logarithme $\ln\colon \mathbb{R}^+ \to \mathbb{R}$ 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.

Proposition 27.67.
Si $z \in \mathbb{C}$ alors

$$\int \frac{1}{x+z} dx = \ln(x+z) \tag{27.288}$$

Démonstration. Il est important de comprendre que la formule (27.288) est un abus de notation pour dire que si nous considérons la fonction

$$\begin{aligned} \varphi\colon \mathbb{R} &\to \mathbb{C} \\ x &\mapsto \ln(x+z) \end{aligned} \tag{27.289}$$

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 φ se décompose en partie réelle et imaginaire qui sont à dériver séparément :

$$\varphi(x) = \ln(|x+z|) + i \arg(x+z). \tag{27.290}$$

Si z est imaginaire pur Nous posons $z = \lambda i$ avec $\lambda \in \mathbb{R}^*$. D'abord nous avons

$$\frac{1}{x+\lambda i} = \frac{x}{x^2+\lambda^2} - i\frac{\lambda}{x^2+\lambda^2}. \tag{27.291}$$

La partie réelle de $\varphi(x)$ est

$$\varphi_1(x) = \ln\left(\sqrt{x^2+\lambda^2}\right), \tag{27.292}$$

dont la dérivée est

$$\varphi_1'(x) = \frac{x}{x^2+\lambda^2}, \tag{27.293}$$

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 (27.251) :

$$\varphi_2'(x) = \frac{-\lambda}{x^2+\lambda}. \tag{27.294}$$

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.

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

$$\varphi'(x) = \varphi_0'(x + s) = \frac{1}{x + s + \lambda i} = \frac{1}{x + z}. \tag{27.295}$$

Tout va bien.

□

Exemple 27.68

Un petit calcul d'intégrale, que nous avions déjà faite dans l'exemple 21.98 (avec la méthode de Rothstein-Trager). En passant par une décomposition en fractions simples :

$$\int \frac{1}{x^3 + x} = \int \left(\frac{1}{x} - \frac{1/2}{x - i} - \frac{1/2}{x + i} \right) \tag{27.296a}$$

$$= \ln(x) - \frac{1}{2}\ln(x - i) - \frac{1}{2}\ln(x + i) \tag{27.296b}$$

$$= \ln(x) - \frac{1}{2}\ln(x^2 + 1). \tag{27.296c}$$

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,

$$\arg(ab) \neq \arg(a) + \arg(b) \tag{27.297}$$

lorsque la somme dépasse les bornes de $]-\pi, \pi]$. Le passage à (27.296c) fonctionne parce que dans le cas particulier des nombres $x+i$ et $x-i$, les arguments se somment à zéro : $\arg(x+i)+\arg(x-i) = 0$.

△

27.8 Théorème de Weierstrass

Théorème 27.69 (Théorème de Weierstrass[406]).
Soit (f_n) une suite de fonctions holomorphes sur un ouvert Ω de $\mathbb{C}$ que nous supposons converger uniformément sur tout compact vers f. Alors f est holomorphe sur Ω et pour tout k nous avons

$$f_n^{(k)} \to f^{(k)} \tag{27.298}$$

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 27.13 :

$$f_n(z) = \frac{1}{2\pi i} \int_{\partial B(a,r)} \frac{f_n(\xi)}{\xi - z} d\xi \tag{27.299}$$

pour tout z dans un boule $B(a, \rho)$ incluse dans $B(a, r)$. Étant donné que le cercle ∂B est compact, elle y est majorée par une constante M. Montrons que de plus nous pouvons choisir M de telle façon

à avoir $|f_n(\xi)| \leqslant M$ pour tout n et tout ξ en même temps. D'abord nous utilisons la continuité de la limite f sur le compact ∂B pour poser $A = \max_{z\in\partial B} |f(z)|$. Ensuite nous considérons un $\epsilon > 0$ et N tel que $\| f_n - f\|_{\partial B} \leqslant \epsilon$ pour tout $n \geqslant N$. Nous savons maintenant que

$$\{|f_n(\xi)| \text{ tel que } n \geqslant N, \xi \in \partial B\} \tag{27.300}$$

est majoré par $A + \epsilon$. Nous posons enfin

$$B = \max_{n\leqslant N} \max_{\xi\in\partial B} |f_n(z)|, \tag{27.301}$$

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 ξ. Nous pouvons donc permuter l'intégrale et la limite sur n :

$$f(z) = \frac{1}{2i\pi} \int_{\partial B} \frac{f(\xi)}{\xi - z}. \tag{27.302}$$

Cela implique que la fonction f est holomorphe par le corollaire 27.16.

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 corollaire 27.29 :

$$f_n^{(k)} = \frac{1}{2\pi i} \int_{\partial B(z_0,r)} \frac{f(\omega)}{(\omega - z)^{k+1}} d\omega. \tag{27.303}$$

Nous voulons la convergence sur tout compact contenu dans l'ouvert Ω. 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 $n \geqslant N_{(z_0,r)}$,

$$\| f_n^{(k)} - f^{(k)}\|_{B(z_0,r)} \leqslant \epsilon. \tag{27.304}$$

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

$$\| f_n^{(k)} - f^{(k)}\|_K \leqslant \epsilon. \tag{27.305}$$

Au travail !

Pour $z \in B(z_0, r)$ nous considérons $r' > r$ tel que $B(z_0, r') \subset \Omega$ et nous avons

$$|f_n^{(k)}(z) - f^{(k)}(z)| = \left| \frac{1}{2\pi i} \int_{\partial B(z_0,r')} \frac{f_n(\xi) - f(\xi)}{(\xi - z)^{k+1}} d\xi \right| \tag{27.306a}$$

$$\leqslant \frac{1}{2\pi} \int_{\partial B(z_0,r')} \frac{|f_n(\xi) - f(\xi)|}{|r - r'|^{k+1}} d\xi. \tag{27.306b}$$

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 $f_n \to f$ uniformément, nous pouvons considérer n assez grand pour que le numérateur soit plus petit que ϵ indépendamment de ξ et de z. Donc pour un n assez grand,

$$|f_n^{(k)}(z) - f^{(k)}(z)| \leqslant \frac{\epsilon}{2\pi} \frac{2\pi r'}{|r - r'|^{k+1}} \tag{27.307}$$

pour tout $z \in B(z_0, r)$. Donc nous avons convergence uniforme $f_n^{(k)} \to f^{(k)}$ sur cette boule. Par l'argument de compacité donné plus haut, nous avons la convergence uniforme sur tout compact. □

Chapitre 28

Analyse fonctionnelle

Théorème 28.1 (Théorème d'isomorphisme de Banach)**.**
Une application linéaire continue et bijective entre deux espaces de Banach est un homéomorphisme.

28.1 Théorème d'Ascoli

Définition 28.2.
*Une partie A d'un espace topologique X est **relativement compacte** dans X si sa fermeture est compacte.*

Proposition 28.3 ([218])**.**
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.

(1) *L'image d'un borné de E par f est relativement compact dans F.*

(2) *L'image par f de la boule unité fermée est relativement compacte dans F.*

(3) *Si (x_n) est une suite bornée dans E, alors nous pouvons en extraire une sous-suite $(x_{\varphi(n)})$ telle que $\big(fx_{\varphi(n)}\big)$ converge dans F.*

Définition 28.4.
*Une application vérifiant les conditions équivalentes de la proposition 28.3 est dite **compacte**.*

Théorème 28.5 (Théorème d'Ascoli[407])**.**
Soient un espace topologique compact K et un espace métrique (E,d). Nous considérons la topologie uniforme sur $C(K,E)$. Une partie A de $C(K,E)$ est relativement compacte si et seulement si les deux conditions suivantes sont remplies :

(1) *A est équicontinu*[1]*,*

(2) *$\forall x \in K$, l'ensemble $\{f(x)$ tel que $f \in A\}$ est relativement compact dans E.*

Corollaire 28.6.
Soient $\big(E,(p_l)\big)$ et $\big(F,(q_k)\big)$ deux espaces localement compacts munis de semi-normes. Nous supposons que E est métrisable et complet. Soit $(T_j)_{j\in\mathbb{N}}$ une suite d'application linéaires $E \to F$ telles que pour tout $x \in E$ il existe $\alpha_x \in F$ tel que

$$T_j x \xrightarrow{F} \alpha_x. \tag{28.1}$$

Si nous posons $Tx = \alpha_x$ alors

(1) *l'application T est linéaire et continue,*

1. Définition 9.69.

(2) pour tout compact K dans E et pour tout k nous avons

$$\lim_{j\to\infty} \sup_{x\in K} q_k(T_j x - Tx) = 0, \tag{28.2}$$

(3) si $x_j \to x$ dans E alors

$$T_j x_j \to Tx \tag{28.3}$$

dans F.

La version suivante du théorème de Banach-Steinhaus est énoncée de façon *ad hoc* pour fonctionner avec l'espace $\mathscr{D}(K)$ des fonctions de classe C^∞ à support dans le compact K. Un énoncé un peu plus fort est donné dans le cadre des espaces de Fréchet dans [117].

Théorème 28.7 (Banach-Steinhaus avec des semi-normes)**.**
Soit (E,d) un espace vectoriel métrique complet dont la topologie est également[2] *donnée par une famille $\mathcal{P}$ de semi-normes. Soit $\{T_\alpha\}_{\alpha\in A}$ une famille d'applications linéaires continues $T_\alpha : E \to \mathbb{R}$ telles que pour tout $x \in E$ nous ayons*

$$\sup_{\alpha\in A} \big|T_\alpha(x)\big| < \infty. \tag{28.4}$$

Alors il existe une constante $C > 0$ et un sous-ensemble fini $J \subset \mathcal{P}$ tels que pour tout $x \in E$ nous ayons

$$\big|T_\alpha(x)\big| \leqslant C \max_{j\in J} p_j(x). \tag{28.5}$$

Démonstration. Pour chaque $k \in \mathbb{N}^*$ nous posons

$$\Omega_k = \{x \in E \text{ tel que } \sup_{\alpha\in A} \big|T_\alpha(x)\big| > k\}. \tag{28.6}$$

Ces ensembles sont des ouverts (pour la même raison que dans la preuve du théorème 12.92) et leur union est E en entier parce que par hypothèse $\sup_{\alpha\in A} \big|T_\alpha(x)\big| < \infty$.

Si les Ω_k étaient tous dense, le théorème de Baire 9.87 nous dit que leur intersection est dense également ; elle est donc non vide et si $x_0 \in \bigcap_{k\in\mathbb{N}} \Omega_k$ nous avons

$$\sup_{\alpha\in A} \big|T_\alpha(x_0)\big| = \infty, \tag{28.7}$$

ce qui contredirait l'hypothèse. Donc les Ω_k ne sont pas tous denses. Soit $k_0 \in \mathbb{N}^*$ tel que Ω_{k_0} n'est pas dense dans E. Il existe donc $x_0 \in E$ et un ouvert autour de x_0 n'intersectant pas Ω_{k_0}.

Nous jouons à présent sur la topologie de E. L'ouvert dont il est question est un d-ouvert et donc un $\mathcal{P}$-ouvert, lequel contient une $\mathcal{P}$-boule ouverte. Cette dernière boule n'est pas spécialement une d-boule, mais c'est un d-ouvert.

Il existe dont J fini dans $\mathcal{P}$ et $\rho > 0$ tels que $B_J(x_0,\rho) \cap \Omega_{k_0} = \varnothing$. Donc pour tout $y \in B_J(x_0,\rho)$ nous avons

$$\sup_{\alpha\in A} \big|T_\alpha(y)\big| \leqslant k_0. \tag{28.8}$$

Si maintenant $y \in B(0,\rho)$, nous avons $y = (y + x_0) - x_0$ et donc

$$\begin{aligned}
\big|T_\alpha(y)\big| &= \big|T_\alpha(y + x_0) - T_\alpha(x_0)\big| && \text{(28.9a)}\\
&\leqslant \big|T_\alpha(y + x_0)\big| + \big|T_\alpha(x_0)\big| && \text{(28.9b)}\\
&\leqslant k_0 + C && \text{(28.9c)}
\end{aligned}$$

où nous avons posé $C = \sup_{\alpha\in A} \big|T_\alpha(x_0)\big|$. En normalisant, sur la boule $B_J(0,1)$ nous avons

$$\big|T_\alpha(y)\big| \leqslant \rho^{-1}(k_0 + C). \tag{28.10}$$

2. Au sens où les ouverts sont les mêmes.

Enfin su $x \in E$ nous avons

$$\frac{x}{\max_{j\in J} p_j(x)} \in B_J(0,1) \tag{28.11}$$

et donc

$$\left|T_\alpha\left(\frac{x}{\max_{j\in J}}p_j(x)\right)\right| \leqslant \rho^{-1}(k_0+C). \tag{28.12}$$

Utilisant encore la linéarité de T_α nous trouvons ce que nous devions trouver :

$$\left|T_\alpha(x)\right| \leqslant \max_{j\in J} p_j(x)\rho^{-1}(k_0+C) \tag{28.13}$$

à redéfinition près de $\rho^{-1}(k_0+C)$ en C. □

28.2 Espaces de Lebesgue L^p

28.2.1 Généralités

Définition 28.8.
Soit $(\Omega,\mathcal{F},\mu)$ un espace mesuré. Deux fonctions à valeurs complexes f et g sur cet espaces sont dites ***équivalentes*** *et nous notons $f\sim g$ si elles sont μ-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 28.9.
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 ϵ 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 28.10.
Pour $p\geqslant 1$, nous introduisons l'opération

$$\|f\|_p = \left(\int_\Omega |f(x)|^p d\mu(x)\right)^{1/p} \tag{28.14}$$

et nous notons $\mathcal{L}^p(\Omega,\mu)$ l'ensemble des fonctions mesurables sur Ω telles que $\|f\|_p<\infty$.

28.11.
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 (28.14).

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 15.141), alors nous considérons un borné mesurable K contenant A et la fonction

$$f(x)=\begin{cases}1 & \text{si } x\in A\\ -1 & \text{si } x\in K\backslash A\\ 0 & \text{sinon.}\end{cases} \tag{28.15}$$

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 $\mathcal{L}^p$.

Lemme 28.12 ([1]).
L'ensemble $\mathcal{L}^p$ est un espace vectoriel sur $\mathbb{C}$.

Démonstration. Pour rappel, nous ne considérons les choses que pour $p \geqslant 1$. 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 $\varphi\colon x \mapsto x^p$ est convexe sur les positifs (lemme 18.85), de telle sorte que

$$\varphi\left(\frac{|a| + |b|}{2}\right) \leqslant \frac{\varphi(|a|) + \varphi(|b|)}{2}, \tag{28.16}$$

ou encore

$$(|a| + |b|)^p \leqslant 2^{p-1}(|a|^p + |b|^p) \tag{28.17}$$

Et d'autre part, nous savons que pour $z_1, z_2 \in \mathbb{C}$, $|z_1 + z_2| \leqslant |z_1| + |z_2|$ (proposition 1.132(5)). Donc

$$\begin{aligned}
\int_\Omega |f + g|^p d\mu &= \int_\Omega |f(\omega) + g(\omega)|^p d\mu(\omega) && \text{(28.18a)}\\
&\leqslant \int_\Omega \big(|f(\omega)| + |g(\omega)|\big)^p d\mu(\omega) && \text{(28.18b)}\\
&\leqslant 2^{p-1}\big(|f(\omega)|^p + |g(\omega)|^p\big)d\mu(\omega) && \text{(28.18c)}\\
&= 2^{p-1}\big(\|f\|_p^p + |g|_p^p\big). && \text{(28.18d)}
\end{aligned}$$

Donc si f et g sont dans $\mathcal{L}^p$, alors $f + g$ est dans $\mathcal{L}^p$. □

28.13.
Il est à noter que nous ne considérons que des valeurs $p \geqslant 1$, précisément parce que la fonction $x \mapsto |x|^p$ n'est pas convexe lorsque $p < 1$.

Dans le même ordre d'idées, si $p \geqslant 1$, alors le $q \in \mathbb{R}$ tel que

$$\frac{1}{p} + \frac{1}{q} = 1 \tag{28.19}$$

est également $q \geqslant 1$. Cela est important pour un certain nombres de théorèmes qui vont venir, en particulier l'inégalité de Hölder (28.59).

Si vous en voulez à propos de $0 < p < 1$, vous pouvez lire [408].

28.14.
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 $\|.\|_p$ s'annule.

28.15.
Soit un espace mesuré non complet[3] $(\Omega, \mathcal{A}, \mu)$. Il existe une partie N de mesure nulle et $A \subset N$ non mesurable. Considérons

$$f(x) = \begin{cases} 2 & \text{si } x \in A \\ 1 & \text{si } x \in N\backslash A \\ 0 & \text{sinon.} \end{cases} \tag{28.20}$$

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 importances pour deux raisons.

— La mesure de Lebesgue est complète par définition (définition 15.127) si nous considérons bien la tribu de Lebesgue et non seulement les boréliens.

3. Définition 15.56.

— 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.

Lemme 28.16.
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 Ω ; son intégrale y vaut zéro. Nous avons

$$\int_\Omega |f(x)|^p d\mu(x) = \int_\Omega \big(|f(x)|^p + h(x)\big) d\mu(x) = \int_\omega |g(x)|^p d\mu(x). \tag{28.21}$$

Cela prouve que la dernière intégrale existe et vaut la même chose que la première. □

Nous pouvons donc considérer la norme $|.|_p$ 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

$$\|[f]\|_p = \|f\|_p, \tag{28.22}$$

formule qui ne dépend pas du représentant par le lemme 28.16.

Proposition 28.17.
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$.

(1) L'ensemble des classes $L^p(\Omega, \mathcal{A}, \mu)$ est un espace vectoriel.

(2) La formule

$$\|[f]\|_p = \left(\int_\Omega |f(x)|^p d\mu(x)\right)^{1/p} \tag{28.23}$$

défini une norme[4] *sur $L^p(\Omega, \mathcal{A}, \mu)$.*

Démonstration. En deux parties.

Espace vectoriel Le lemme 28.12 dit que $\mathcal{L}^p$ est un espace vectoriel. Une structure d'espace vectoriel est donnée en posant

$$[f] + [g] = [f + g] \tag{28.24}$$

et

$$\lambda[f] = [\lambda f] \tag{28.25}$$

qui sont deux définitions correctes parce qu'elles ne dépendent pas du choix du représentant. De plus le lemme 28.12 dit que $f + g \in \mathcal{L}^p$ dès que $f, g \in \mathcal{L}^p$. Donc $[f + g] \in L^p$ dès que $[f], [g] \in Ł^p$.

Norme Pour être une norme, il faut vérifier les trois propriétés de la définition 7.106.

D'abord, si $\|[f]\|_p = 0$, nous avons

$$\int_\Omega |f(x)|^p d\mu(x) = 0, \tag{28.26}$$

ce qui par le lemme 15.179 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.

Ensuite pour $\lambda \in \mathbb{C}$ et $f \in L^p$ nous avons

$$\|\lambda f\|_p = \left(\int_\Omega |\lambda f(\omega)|^p d\mu(\omega)\right)^{1/p} \tag{28.27a}$$

$$= \left(\int_\Omega |\lambda|^p |f(\omega)|^p d\mu(\omega)\right)^{1/p} \tag{28.27b}$$

$$= |\lambda| \|f\|_p. \tag{28.27c}$$

4. Définition 7.106.

Et enfin, en suivant le calcul (28.18) nous avons

$$\|f+g\|_p \leqslant \|f\|_p + \|g\|_p. \tag{28.28}$$

□

28.18.
À partir de maintenant $\big(L^p(\Omega,\mu), \|.\|_p\big)$ 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.

Proposition 28.19 ([409]).
Soit $1 \leqslant p \leqslant \infty$ 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 μ-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 $1 \leqslant p < \infty$. Nous considérons une sous-suite $[h_n]$ de $[f_n]$ telle que

$$\|[h_j]-[f]\|_p < 2^{-j}, \tag{28.29}$$

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

$$\int_\Omega u_k d\mu = \int_\omega |h_k(x)-f(x)|^p d\mu(x) = \|h_k-f\|_p^p \leqslant 2^{-kp}. \tag{28.30}$$

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 15.160 :

$$\int_\Omega \left(\sum_{k=0}^{\infty} u_k(x)\right) d\mu(x) = \sum_{k=0}^{\infty}\int_\Omega u_k(x)d\mu(x) \leqslant \sum_{k=0}^{\infty} 2^{-kp} < \infty. \tag{28.31}$$

Le fait que l'intégrale de la fonction $\sum_k u_k$ est finie implique que cette fonction est finie μ-presque partout. Donc le terme général tend vers zéro presque partout, c'est-à-dire

$$|h_k(x)-f(x)|^p \to 0. \tag{28.32}$$

Cela signifie que $h_k \to f$ 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.

Exemple 28.20
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. Cet exemple provient de bibmath.net.

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 $f_n \to 0$ 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.

Proposition 28.21.
Soit $f \in L^p(\Omega)$ avec $1 \leqslant p < \infty$ et (f_n) une suite de fonctions convergeant ponctuellement vers f et telle que $|f_n| \leqslant |f|$. Alors $f_n \xrightarrow{L^p} f$.

Démonstration. Nous avons immédiatement $|f_n(x)|^p \leqslant |f(x)|^p$, de telle sorte que le théorème de la convergence dominée implique que $f_n \in L^p$. La convergence dominée donne aussi que $\|f_n\|_p \to \|f\|_p$, mais cela ne nous intéresse pas ici.

Nous posons $h_n(x) = |f_n(x) - f(x)|$. En reprenant la formule de majoration (28.17) et en tenant compte du fait que $|f_n(x)| \leqslant |f(x)|$, nous avons

$$h_n(x) \leqslant 2^{p-1}\big(|f_n(x)|^p + |f(x)|^p\big) \leqslant 2^p |f(x)|^p, \tag{28.33}$$

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 15.184) :

$$\lim_{n\to\infty} \|f_n - f\|_p^p = \lim_{n\to\infty} \int_{\mathbb{R}^d} |f_n(x) - f(x)|^p dx = \int_{\mathbb{R}^d} \lim_{n\to\infty} h_n(x) dx = 0. \tag{28.34}$$

□

Proposition 28.22 ([1])**.**
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et une fonction mesurable $f : \Omega \to \mathbb{C}$ telle que

$$\int_\Omega f \mathbb{1}_A d\mu = 0 \tag{28.35}$$

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 φ est intégrable signifie que $\mathrm{Re}(f)^+$, $\mathrm{Re}(f)^-$, $\mathrm{Im}(f)^+$ et $\mathrm{Im}(f)^-$ sont intégrables.

Les parties

$$\{x \in \Omega \text{ tel que } \mathrm{Re}(f)^+(x) \in [k, k+1[\}_{k=1,\dots} \tag{28.36}$$

et

$$\{x \in \Omega \text{ tel que } \mathrm{Re}(f)^+(x) \in [\frac{1}{k+1}, \frac{1}{k}[\}_{k=1,\dots} \tag{28.37}$$

sont mesurables en tant qu'images inverses de mesurables par la fonction mesurable $\mathrm{Re}(f)^+$.

Nous notons $\{A_i\}_{i\in\mathbb{N}}$ une énumération quelconque de ces parties. L'important est que

$$\bigcup_{i\in\mathbb{N}} A_i = \{x \in \Omega \text{ tel que } \mathrm{Re}(f)^+(x) \neq 0\} \tag{28.38}$$

et que pour chaque i, il existe $\alpha_i > 0$ tel que $\mathrm{Re}(f)^+(x) > \alpha_i$ pour tout x dans A_i.

Par hypothèse nous avons $\int_\Omega f\mathbb{1}_{A_i} d\mu = 0$. En particulier,

$$0 = \int_\Omega \mathrm{Re}(f\mathbb{1}_{A_i})^+ = \int_\Omega \mathrm{Re}(f)^+(x)\mathbb{1}_{A_i}(x) d\mu(x). \tag{28.39}$$

Mais pour tout $x \in A_i$ nous avons $\mathrm{Re}(f)^+(x) > \alpha_i > 0$, donc

$$0 = \int_\Omega \mathrm{Re}(f)^+(x)\mathbb{1}_{A_i}(x) d\mu(x) \geqslant \alpha_i \int_\Omega \mathbb{1}_{A_i} = \alpha_i \mu(A_i) \geqslant 0. \tag{28.40}$$

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 $\mathrm{Re}(f)^+$ est non nulle seulement sur une partie de mesure nulle.

Le même raisonnement pour $\mathrm{Re}(f)^-$, $\mathrm{Im}(f)^+$ et $\mathrm{Im}(f)^-$ donne que f est non nulle sur une partie de mesure nulle. Donc $f = 0$ au sens des classes dans L^p. □

28.2.2 L'espace L^∞

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 que point. Nous cherchons alors à définir une notion de supremum qui ne tient pas compte des ensembles de mesure nulle.

Définition 28.23.
*Soit $f : \Omega \to \mathbb{C}$. Un nombre M est un **majorant essentiel** de f si*

$$\mu\big(|f(x)| \leqslant M\big) = 0. \tag{28.41}$$

Nous posons alors

$$N_\infty(f) = \inf\{M \text{ tel que } |f(x)| \leqslant M \text{ presque partout}\}. \tag{28.42}$$

Cela revient à prendre le supremum à ensemble de mesure nulle près.

Définition 28.24.
Nous définissons alors les espaces de Lebesgue correspondants :

$$\mathcal{L}^\infty(\Omega) = \{f : \Omega \to \mathbb{C} \text{ tel que } N_\infty(f) < \infty\}, \tag{28.43}$$

et L^∞ en est le quotient usuel.

28.25.
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^∞, 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 ∞, alors que les éléments de L^∞ 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^∞.
(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 intolérable :

(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 $\|[f]\|_{L^\infty}$.

28.26.
Tout ceci pour dire que si f_k sont des fonctions et si f est une fonction, nous avons la convergence

$$f_k \xrightarrow{L^\infty} f \tag{28.44}$$

si et seulement si $N_\infty(f_k - f) \to 0$. Cette convergence (qui se sert des abus tolérables de notations) est équivalente à la convergence

$$[f_k] \xrightarrow{L^\infty} [f]. \tag{28.45}$$

Mais ce n'est pas du tout équivalent à $\|f_k - f\|_\infty \to 0$. Tout au plus, il est vrai que si $\alpha \in L^\infty$ (donc α est une classe), il existe un représentant $f \in \alpha$ tel que $\|f\|_\infty = \alpha_{L^\infty}$.

Lemme 28.27 ([1]).
Si f_k est une suite de fonctions telle que $\|f_k - f\|_\infty \to 0$, alors

$$f_k \xrightarrow{L^\infty} f \tag{28.46}$$

où la convergence signifie $N_\infty(f_k - f) \to 0$.

Démonstration. Si g est une fonction nous avons toujours $N_\infty(g) \leqslant \|g\|_\infty$. Donc

$$N_\infty(f_k - f) \leqslant \|f_k - f\|_\infty \to 0. \tag{28.47}$$

□

28.28.
Attention toutefois que ce lemme ne signifie pas que si $\|f_k - f\|_\infty \to 0$, 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^∞.

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^∞.

Exemple 28.29
Même si g est bornée, nous n'avons pas spécialement $\|g\|_\infty = N_\infty(g)$. Par exemple

$$\begin{aligned} g\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{sinon .} \end{cases} \end{aligned} \tag{28.48}$$

Cette fonction vérifie $\|g\|_\infty = 1$ mais $N_\infty(g) = 0$. △

28.2.3 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 28.30.
Nous avons les égalités suivantes d'espaces

$$L^p\big(]0, 2\pi[\big) = L^p\big([0, 2\pi]\big) = L^p(S^1) \tag{28.49}$$

au sens où il existe des bijections isométriques de l'un à l'autre. Ici nous sous-entendons la mesure de Lebesgue partout[5].

Démonstration. Voici une application bien définie où le crochet dénote la prise de classe :

$$\begin{aligned} \psi\colon L^p\big(]0, 2\pi[\big) &\to L^p\big([0, 2\pi]\big) \\ [f] &\mapsto \text{la classe de } f_e(x) = \begin{cases} f(x) & \text{si } x \in]0, 2\pi[\\ 0 & \text{si } x = 0 \text{ ou } x = 2\pi. \end{cases} \end{aligned} \tag{28.50}$$

Injective Si $[f] = [g]$ dans $L^p\big(]0, 2\pi[\big)$ 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 ψ sont donc égales dans $L^p\big([0, 2\pi]\big)$.

5. 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.

Surjective Un élément de $L^p\big([0,2\pi]\big)$ est l'image de sa restriction ... ou plutôt l'image de la classe de la restriction d'un quelconque de ses représentants.

Isométrie L'intégrale qui donne la norme sur L^p ne change pas selon que nous ajoutions ou non les bornes au domaine d'intégration.

De la même manière nous avons

$$L^p\big([0,2\pi[\big) = L^p\big([0,2\pi]\big). \tag{28.51}$$

En ce qui concerne l'identification avec $L^p(S^1)$, il faut passer par l'isométrie $\varphi\colon [0,2\pi[\to S^1$ 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

$$\int_0^{2\pi} f(t)dt = \int_{S^1} (f\circ\varphi^{-1})(z)dz. \tag{28.52}$$

□

28.2.4 Inégalité de Jensen, Hölder et de Minkowski

Proposition 28.31 (Inégalité de Jensen[263])**.**
Soit un espace mesuré de probabilité[6] $(\Omega,\mathcal{A},\mu)$ *ainsi qu'une fonction convexe* $f\colon\mathbb{R}\to\mathbb{R}$ *et une application* $\alpha\colon\Omega\to\mathbb{R}$ *tels que* α *et* $f\circ\alpha$ *soient intégrables sur* Ω*. Alors*

$$f\Big(\int_\Omega \alpha\,d\mu\Big) \leqslant \int_\Omega (f\circ\alpha)d\mu. \tag{28.53}$$

Démonstration. Soient $a\in\mathbb{R}$ et c_a le nombre donné par la proposition 18.88 : pour tout $\omega\in\Omega$ nous avons

$$f\big(\alpha(\omega)\big) - f(a) \geqslant c_a\big(\alpha(\omega)-a\big). \tag{28.54}$$

Cela est en particulier vrai pour $a = \int_\Omega \alpha\,d\mu$. Nous intégrons l'inégalité (28.54) sur Ω en nous souvenant que $\int d\mu = 1$:

$$\int_\Omega (f\circ\alpha)d\mu - \int_\Omega f(a)d\mu \geqslant c_a\Big(\int_\Omega \alpha - \int_\Omega a\Big) \tag{28.55a}$$

$$\int_\Omega (f\circ\alpha)d\mu - f(a) \geqslant c_a(a-a) \tag{28.55b}$$

$$f(a) \leqslant \int_\Omega (f\circ\alpha)d\mu. \tag{28.55c}$$

Cette dernière inégalité est celle que nous devions prouver. □

Corollaire 28.32.
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* $1\leqslant p<+\infty$*. Alors*

$$\Big|\int_\Omega \alpha(s)d\mu(s)\Big| \leqslant \|\alpha\|_p. \tag{28.56}$$

Démonstration. Il suffi d'utiliser l'inégalité de Jensen sur la fonction convexe $f(x) = |x|^p$. Nous avons alors

$$\Big|\int_\Omega \alpha(s)d\mu(s)\Big|^p \leqslant \int_\Omega |\alpha(s)|^p d\mu(s), \tag{28.57}$$

c'est-à-dire

$$\Big|\int_\Omega \alpha(s)d\mu(s)\Big| \leqslant \Big[\int_\Omega |\alpha(s)|^p d\mu(s)\Big]^{1/p} = \|\alpha\|_p \tag{28.58}$$

où ma norme $\|.\|_p$ est prise au sens de la mesure μ. □

6. C'est-à-dire que $\int_\Omega d\mu = 1$.

Proposition 28.33 (Inégalité de Hölder)**.**
Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $1 \leqslant p, q \leqslant \infty$ satisfaisant $\frac{1}{p} + \frac{1}{q} = 1$. Si $f \in L^p(\Omega)$, $g \in L^q(\Omega)$, alors le produit fg est dans $L^1(\Omega)$ et nous avons

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q. \tag{28.59}$$

Plus généralement si $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$ alors

$$\|fg\|_r \leqslant \|f\|_p \|g\|_q \tag{28.60}$$

Ce qui est appelé « inégalité de Hölder » est généralement l'inéquation (28.59).

Démonstration. Nous allons prouver l'inégalité (28.60). D'abord nous supposons $\|g\|_q = 1$ et nous posons

$$A = \{x \in \Omega \text{ tel que } |g(x)| > 0\}. \tag{28.61}$$

Hors de A, les intégrales que nous allons écrire sont nulles. Nous avons

$$\|fg\|_r^p = \Big| \int_A |f|^r |g|^{r-q} |g|^q \Big|^{p/r}, \tag{28.62}$$

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 [7] parce que $p/r > 1$, ce qui fait de $x \mapsto |x|^{p/r}$ une fonction convexe. Nous avons alors

$$\|fg\|_r^p \leqslant \int_A \big(|f|^r |g|^{r-q}\big)^{p/r} |g|^q d\mu \tag{28.63a}$$

$$= \int_A |f|^p |g|^{p(r-q)/r} |g|^q d\mu \tag{28.63b}$$

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

$$\|fg\|_r^p \leqslant \int_A |f|^p d\mu \leqslant \|f\|_p^p. \tag{28.64}$$

La dernière inégalité est le fait que le domaine A n'est pas tout le domaine Ω.

Si maintenant $\|g\|_q \neq 1$ alors nous calculons

$$\|fg\|_r = \|g\|_q \|f \frac{g}{\|g\|_q}\|_r \leqslant \|g\|_q \|f\|_p \tag{28.65}$$

en appliquant la première partie à la fonction $\frac{g}{\|g\|_q}$ qui est de norme 1. □

Remarque 28.34.
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

$$\|f\|_1 \leqslant \|f\|_2. \tag{28.66}$$

Lemme 28.35.
Lorsque I est borné nous avons $L^2(I) \subset L^1(I)$. 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 28.33 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. □

7. Proposition 28.31.

Corollaire 28.36 ([410]).
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 :

$$\lim_{n\to\infty} \int_I u_n = \int_I u. \tag{28.67}$$

Démonstration. Nous considérons la forme linéaire

$$\begin{aligned} T\colon L^2(I) &\to \mathbb{R} \\ u &\mapsto \int_I u. \end{aligned} \tag{28.68}$$

Elle est bien définie par l'inégalité de Hölder $\|fg\|_1 \leqslant \|f\|_2\|g\|_2$ appliqué à $g(x) = 1$ qui vérifie $\|g\|_2 = 1$. Nous avons aussi

$$T(u) \leqslant \int_I |u| \leqslant \|u\|_1 \leqslant \|u\|_2 \tag{28.69}$$

où la dernière inégalité est celle de Hölder 28.33. 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. □

Proposition 28.37 (Inégalité de Minkowski[411, 412]).
Si $1 \leqslant p < \infty$ et si $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ alors

(1) $\|f + g\|_p \leqslant \|f\|_p + \|g\|_p$

(2) Il y a égalité si et seulement si les vecteurs $f(x)$ et $g(x)$ sont presque partout colinéaires : il existe α, β tels que $\alpha f + \beta g = 0$ presque partout.

(3) Si $f(x,y)$ est mesurable sur l'espace produit $\big(X \times Y, \mu \otimes \nu\big)$ et si $p \geqslant 1$, alors

$$\left\| x \mapsto \int_Y f(x,y) d\nu(y) \right\|_p \leqslant \int_Y \|f_y\|_p d\nu(y) \tag{28.70}$$

où $f_y(x) = f(x,y)$.

La partie (3) est une généralisation de l'inégalité triangulaire (c'est-à-dire du point (1)) dans le cas où nous n'avons pas une somme de deux fonctions mais d'une infinité paramétrée par $y \in Y$. Elle sera le plus souvent utilisée sous la forme déballée :

$$\left[\int_X \Big(\int_Y |f(x,y)| d\nu(y)\Big)^p d\mu(x)\right]^{1/p} \leqslant \int_Y \Big(\int_X |f(x,y)|^p d\mu(x)\Big)^{1/p} d\nu(y). \tag{28.71}$$

28.2.5 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 28.38 que de telles inclusions et inégalités sont possibles pour $L^p\big([0,1]\big)$.

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 15.248.

$\underline{L^p \not\subseteq L^q}$ La fonction

$$f(x) = \begin{cases} \frac{1}{x^{1/q}} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases} \tag{28.72}$$

est dans L^p mais pas dans L^q. En effet

$$\|f\|_p^p = \int_0^1 \frac{1}{x^{p/q}} dx < \infty \tag{28.73}$$

parce que $p < q$ et $p/q < 1$. Par contre

$$\|f\|_q^q = \int_0^1 \frac{1}{x} dx = \infty. \tag{28.74}$$

$\underline{L^q \not\subseteq L^p}$ La fonction

$$f(x) = \begin{cases} \frac{1}{x^{1/p}} & \text{si } x > 1 \\ 0 & \text{sinon} \end{cases} \tag{28.75}$$

est dans L^q mais pas dans L^p. En effet

$$\|f\|_p^p = \int_1^\infty \frac{1}{x} = \infty \tag{28.76}$$

alors que

$$\|f\|_q^q = \int_1^\infty \frac{1}{x^{q/p}} dx < \infty. \tag{28.77}$$

$\underline{\text{Exemple de } \|f\|_p > \|f\|_q}$ La fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in [0,2] \\ 0 & \text{sinon.} \end{cases} \tag{28.78}$$

Nous avons

$$\|f\|_p = 2^{1/p} \tag{28.79a}$$
$$\|f\|_q = 2^{1/q}. \tag{28.79b}$$

Mais comme $p < q$ donc $\|f\|_p > \|f\|_q$.

$\underline{\text{Exemple de } \|f\|_p < \|f\|_q}$ La fonction

$$f(x) = \begin{cases} 1 & \text{si } x \in [0, \frac{1}{2}] \\ 0 & \text{sinon.} \end{cases} \tag{28.80}$$

Alors

$$\|f\|_p = \frac{1}{2^{1/p}} \tag{28.81a}$$
$$\|f\|_q = \frac{1}{2^{1/q}} \tag{28.81b}$$

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.

Proposition 28.38 ([6])**.**
Inclusions et inégalités dans le cas d'un ensemble de mesure finie.

(1) Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré fini et $1 \leqslant p \leqslant +\infty$. Alors $L^q(\Omega) \subset L^p(\Omega)$ dès que $p \leqslant q$.

(2) Si $1 < p < 2$ et si $f \in L^2\big([0,1]\big)$ alors $\|f\|_p \leqslant \|f\|_2$.

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

$$A = \{x \in [0,1] \text{ tel que } |f(x)| \geqslant 1\}. \tag{28.82}$$

Étant donné que $p \leqslant q$ nous avons $|f|^p \leqslant |f|^q$ 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 Ω) et sur A^c nous avons $|f(x)| \leqslant 1$ et donc $|f|^p \leqslant 1$. 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 (28.59) avec ces fonctions. D'une part

$$\|f\|_1 = \||f|^p\|_1 = \|f\|_p^p. \tag{28.83}$$

D'autre part

$$\||f|^p\|_{2/p} = \left(\int |f|^2 \right)^{p/2} = \|f\|_2^p. \tag{28.84}$$

Donc $\|f\|_p^p \leqslant \|f\|_2^p$, ce qui prouve l'assertion (2) parce que $p > 1$. □

Remarque 28.39.
Nous n'avons cependant pas $L^2\big([0,1]\big) = L^p\big([0,1]\big)$ parce que l'exemple (28.72) fonctionne encore :

$$f(x) = \frac{1}{\sqrt{x}} \tag{28.85}$$

pour $x \in [0,1]$ donne bien

$$\|f\|_2 = \int_0^1 \frac{1}{x} = \infty \tag{28.86}$$

et $\|f\|_p = \int_0^1 \frac{1}{x^{p/2}} < \infty$ parce que $1 < p < 2$.

28.2.6 Complétude

Théorème 28.40 ([413, 414])**.**
Pour $1 \leqslant p < \infty$, 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 $\|f_p - f_q\|_p \leqslant 2^{-i}$ pour tout $p, q \geqslant N_i$. Nous considérons la sous-suite $g_i = f_{N_i}$, de telle sorte qu'en particulier

$$\|g_i - g_{i-1}\|_p \leqslant 2^{-i}. \tag{28.87}$$

Pour chaque j nous considérons la somme télescopique

$$g_j = g_0 + \sum_{i=1}^{j} (g_i - g_{i-1}) \tag{28.88}$$

et l'inégalité

$$|g_j| \leqslant |g_0| + \sum_{i=1}^{j} |g_i - g_{i-1}|. \tag{28.89}$$

Nous allons noter

$$h_j = |g_0| + \sum_{i=1}^{j} |g_i - g_{i-1}|. \tag{28.90}$$

La suite de fonctions (h_j) ainsi définie est une suite croissante de fonctions positive 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

$$\lim_{j\to\infty} h_j^p = h^p, \tag{28.91}$$

puis par le théorème de la convergence monotone (théorème 15.160) nous avons

$$\lim_{j\to\infty} \int_\Omega h_j^p d\mu = \int_\Omega h^p d\mu. \tag{28.92}$$

Utilisant à présent la continuité de la fonction $x \mapsto x^{1/p}$ nous trouvons

$$\lim_{j\to\infty} \left(\int h_j^p \right)^{1/p} = \left(\int |h|^p \right)^{1/p}. \tag{28.93}$$

Nous avons donc déjà montré que

$$\lim_{j\to\infty} \|h_j\|_p = \left(\int |h|^p \right)^{1/p} \tag{28.94}$$

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 28.37) et l'inégalité (28.87) nous trouvons

$$\|h_j\|_p \leqslant \|g_0\|_p + \sum_{i=1}^{j} \|g_i - g_{i-1}\|_p \leqslant \|g_0\|_p + 1. \tag{28.95}$$

En passant à la limite,

$$\left(\int |h|^p \right)^{1/p} = \lim_{j\to\infty} \|h_j\|_p \leqslant \|g_0\|_p + 1 < \infty. \tag{28.96}$$

Par conséquent $\int |h|^p$ est finie et

$$h \in L^p(\Omega, \mathcal{A}, \mu). \tag{28.97}$$

En particulier, l'intégrale $\int h$ est finie (parce que $p \geqslant 1$) et donc que $h(x) < \infty$ pour presque tout $x \in \Omega$.

Nous savons que $h(x)$ est la limite des sommes partielles (28.90), en particulier la série

$$\sum_{j=1}^{\infty} |g_i - g_{i-1}| \tag{28.98}$$

converge ponctuellement. En vertu du corollaire 15.186, la série de terme général $g_i - g_{i-1}$ converge ponctuellement. La suite g_i converge donc vers une fonction que nous notons g. Par ailleurs la suite g_i est dominée par $h \in L^p$, le théorème de la convergence dominée (théorème 15.184) implique que

$$\lim_{j\to\infty} \|g_j - g\|_p = 0. \tag{28.99}$$

Nous allons maintenant prouver que $\lim_{n\to\infty} \|f_n - g\|_p = 0$. Soit $\epsilon > 0$. Pour tout n et i nous avons

$$\|f_n - g\|_p = \|f_n - f_{N_i} + f_{N_i} - g\|_p \leqslant \|f_n - f_{N_i}\|_p + \|f_{N_i} - g\|_p. \tag{28.100}$$

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 (28.97)) que $h \in L^p$, mais $|g_i| \leqslant h^p$, 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. □

Théorème 28.41 (Fischer-Riesz[43]).
Soit un ouvert Ω de $\mathbb{R}^n$ et $p \in [1, \infty]$. Alors

(1) Toute suite convergente dans $L^p(\Omega)$ admet une sous-suite convergente presque partout sur Ω.

(2) La sous-suite donnée en (1) est dominée par un élément de $L^p(\Omega)$.

(3) L'espace $L^p(\Omega)$ 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.

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 $k \geqslant 1$, il existe $N_k \geqslant 0$ tel que si $m, n \geqslant N_k$, on a

$$\|f_m - f_n\|_\infty \leqslant \frac{1}{k}. \tag{28.101}$$

En particulier, il existe un ensemble de mesure nulle E_k sur lequel

$$|f_m(x) - f_n(x)| \leqslant \frac{1}{k}. \tag{28.102}$$

D'après le lemme 15.27, 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 $m, n \geqslant N_k$, alors

$$|f_n(x) - f_m(x)| \leqslant \frac{1}{k} \tag{28.103}$$

pour tout x hors de E. Donc pour chaque $x \in \Omega\backslash E$, la suite $n \mapsto f_n(x)$ est de Cauchy dans $\mathbb{R}$ et converge donc. Cela défini donc une fonction

$$\begin{aligned} f\colon \Omega\backslash E &\to \mathbb{R} \\ x &\mapsto \lim_{n\to\infty} f_n(x). \end{aligned} \tag{28.104}$$

Cela prouve le point (1) : la convergence ponctuelle.
En passant à la limite $n \to \infty$ dans l'équation 28.103 et tenant compte que cette majoration tient pour presque tout x dans Ω, nous trouvons

$$\|f - f_n\|_\infty \leqslant \frac{1}{k}. \tag{28.105}$$

Donc non seulement f est dans L^∞, mais en plus la suite (f_n) converge vers f au sens L^∞, 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^p.

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 $\varphi\colon \mathbb{N} \to \mathbb{N}$ une fonction strictement croissante telle que pour tout $n \geqslant 1$ nous ayons

$$\|f_{\varphi(n+1)} - f_{\varphi(n)}\|_p \leqslant \frac{1}{2^n}. \tag{28.106}$$

Pour créer la fonction φ, 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 par $\varphi(1) = N_1$ et $\varphi(n+1) > \max\{N_n, \varphi(n-1)\}$. Ensuite nous considérons la fonction

$$g_n(x) = \sum_{k=1}^{n} |f_{\varphi(k+1)}(x) - f_{\varphi(k)}(x)|. \tag{28.107}$$

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 :

$$\|g_n\|_p \leqslant \sum_{k=1}^{n} \|f_{\varphi(k+1)} - f_{\varphi(k)}\|_p \leqslant \sum_{k=1}^{n} \frac{1}{2^k} \leqslant \sum_{k=1}^{\infty} \frac{1}{2^k} = 1. \tag{28.108}$$

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 15.160 sur la convergence monotone. Il existe donc une fonction $g \in L^p(\Omega)$ telle que $g_n \to g$ presque partout.

Soit un $x \in \Omega$ pour lequel $g_n(x) \to g(x)$; alors pour tout $n \geqslant 2$ et $\forall q \geqslant 0$,

$$|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| = \left| f_{\varphi(n+q)}(x) - \sum_{k=1}^{q-1} f_{\varphi(n+k)}(x) - \sum_{k=1}^{q-1} f_{\varphi(n+k)}(x) - f_{\varphi(n)}(x) \right| \tag{28.109a}$$

$$= \left| \sum_{k=1}^{q} f_{\varphi(n+k)} - \sum_{k=1}^{q} f_{\varphi(n+k-1)}(x) \right| \tag{28.109b}$$

$$\leqslant \sum_{k=1}^{q} \left| f_{\varphi(n+k)}(x) - f_{\varphi(n+k-1)}(x) \right| \tag{28.109c}$$

$$= g_{n+q+1}(x) - g_{n+1}(x) \tag{28.109d}$$

$$\leqslant g(x) - g_{n-1}(x). \tag{28.109e}$$

Nous prenons la limite $n \to \infty$; la dernière expression tend vers zéro et donc

$$|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)| \to 0 \tag{28.110}$$

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

$$\begin{aligned} f \colon \Omega \backslash E &\to \mathbb{R} \\ x &\mapsto \lim_{n\to\infty} f_{\varphi(n)}(x) \end{aligned} \tag{28.111}$$

où E est de mesure nulle. Montrons que f est bien dans $L^p(\Omega)$; pour cela nous complétons la série d'inégalités (28.109) en

$$\big|f_{\varphi(n+q)}(x) - f_{\varphi(n)}(x)\big| \leqslant g(x) - g_{n-1}(x) \leqslant g(x). \tag{28.112}$$

En prenant la limite $q \to \infty$ nous avons l'inégalité

$$|f(x) - f_{\varphi(n)}(x)| \leqslant g(x) \tag{28.113}$$

pour presque tout $x \in \Omega$, c'est-à-dire pour tout $x \in \Omega \backslash E$. Cette inégalité implique deux choses valables pour presque tout x dans Ω :

$$f(x) \in B\big(g(x), f_{\varphi(n)}(x)\big) \tag{28.114a}$$

$$f_{\varphi(n)}(x) \leqslant |f(x)| + |g(x)|. \tag{28.114b}$$

La première inégalité assure que $|f|^p$ est intégrable sur $\Omega \backslash E$ 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

$$\|f - f_{\varphi(n)}\|_p \to 0. \tag{28.115}$$

L'inégalité (28.113) nous donne aussi, toujours pour presque tout $x \in \Omega$:

$$\big|f(x) - f_{\varphi(n)}(x)\big|^p \leqslant g(x)^p \tag{28.116}$$

ce qui signifie que la suite [8] $|f - f_{\varphi(n)}|^p$ est dominée par la fonction $|g|^p$ qui est intégrable sur $\Omega \backslash E$ et tout autant sur Ω parce que E est négligeable ; cela prouve au passage le point (2), et le théorème de la convergence dominée de Lebesgue (15.184) nous dit que

$$\lim_{n\to\infty} \int_\Omega \big|f(x) - f_{\varphi(n)}(x)\big|^p dx = \int_\Omega \lim_{n\to\infty} \big|f(x) - f_{\varphi(n)}(x)\big| dx = 0. \tag{28.117}$$

8. À ce point, [43] 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 (28.115).

Cette dernière suite d'égalités se lit de la façon suivante :

$$\lim_{n\to\infty} \|f - f_{\varphi(n)}\|_p = \Big\| \lim_{n\to\infty} |f - f_{\varphi(n)}| \Big\|_p = 0. \tag{28.118}$$

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.

□

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 Weiner, définition 45.1.

Théorème 28.42 (Théorème de Weinersmith[415, 416])**.**
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 $A, B \subset L^p(\Omega, \mathcal{A}, \mu)$ telles que $A \cap B = \varnothing$ 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 11.18, c'est à dire

$$\|f + g\|_p^2 + \|f - g\|_p^2 = 2\|f\|_p^2 + 2\|g\|_p^2. \tag{28.119}$$

Pour $1 \leqslant p < \infty$ Soient donc A, B comme dans l'hypothèse. Nous considérons les fonctions

$$f = \frac{1}{\mu(A)^{1/p}} \mathbb{1}_A \tag{28.120a}$$

$$g = \frac{1}{\mu(B)^{1/p}} \mathbb{1}_B. \tag{28.120b}$$

En ce qui concerne les normes L^p de f et g, c'est un calcul simple :

$$\|f\|_p^2 = \left(\int_\Omega |f(\omega)|^p d\mu \right)^{2/p} = \left(\int_A \frac{1}{\mu(A)} d\mu \right)^{2/p} = 1. \tag{28.121}$$

De même pour g : $\|f\|_p^2 = \|g\|_p^2 = 1$. Donc

$$2\|f\|_p^2 + 2\|g\|_p^2 = 4 \tag{28.122}$$

En ce qui concerne la somme,

$$\|f + g\|_p^2 = \left(\int_A \frac{d\mu}{\mu(A)} + \int_B \frac{1}{\mu(B)} \right)^{2/p} = 2^{2/p}. \tag{28.123}$$

Pour la différence, la seule subtilité à voir est que

$$\int_\Omega |\mathbb{1}_A - \mathbb{1}_B|^p = \int_A |\mathbb{1}_A - \mathbb{1}_B|^p + \int_B |\mathbb{1}_A - \mathbb{1}_B| = \int_A |\mathbb{1}_A| + \int_B |-\mathbb{1}_B| = \int_A \mathbb{1}_A + \int_B \mathbb{1}_B. \tag{28.124}$$

Ce n'est pas de la magie que le moins se change en plus. Bref, pour la différence nous avons

$$\|f-g\|^2 = \left(\int_\Omega |f(\omega)-g(\omega)|^p d\mu\right)^{2/p} \tag{28.125a}$$
$$= \left(\int_\Omega |\frac{1}{\mu(A)^{1/p}}\mathbb{1}_A(\omega) - \frac{1}{\mu(B)^{1/p}}\mathbb{1}_B(\omega)|^p d\mu\right)^{2/p} \tag{28.125b}$$
$$= \left(\int_A |\frac{\mathbb{1}_A(\omega)}{\mu(A)^{1/p}}|^p + \int_B |-\frac{\mathbb{1}_B(\omega)}{\mu(B)^{1/p}}|^p\right)^{2/p} \tag{28.125c}$$
$$= 2^{2/p}. \tag{28.125d}$$

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

$$4 = 2\times 2^{2/p}. \tag{28.126}$$

Cela est uniquement valable pour $p=2$.

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 (28.119) 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 28.43([416])
Soit la mesure de Dirac sur $\mathbb{R}$, c'est à dire

$$\delta(A) = \begin{cases} 1 & \text{si } 0\in A \\ 0 & \text{sinon.} \end{cases} \tag{28.127}$$

Nous allons prouver que pour tout $p\in[0,\infty]$, l'espace $L^p(\mathbb{R},\mathcal{B}\text{or}(\mathbb{R}),\delta)$ est un espace de Hilbert. Pour cela nous introduisons le produit hermitien

$$\langle f,g\rangle = f(0)\overline{g(0)}. \tag{28.128}$$

Nous avons par ailleurs la norme

$$\|f\|^p = \int_{\mathbb{R}} |f(x)|^p d\delta = f(0)^p. \tag{28.129}$$

Donc oui, $\|f\|_p = \sqrt{\langle f,f\rangle}$. △

28.2.7 Théorèmes d'approximation

Proposition 28.44 ([250], thème 2)**.**
Soient $1\leqslant p\leqslant\infty$ et un espace mesuré $(\Omega,\mathcal{A},\mu)$. Alors les fonctions étagées[9] *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 $f:\Omega\to[0,\infty[$. 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.

9. Définition 15.100. Pour rappel, une fonction est simple lorsqu'elle prend un nombre fini de valeurs, et elle est étagée lorsqu'elle est en outre mesurable.

Par le théorème 15.105, il existe une suite croissante de fonctions étagées $\phi_n : \Omega \to [0,\infty[$ telles que $\phi_n \to f$ 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 28.10. Notre devoir est maintenant de prouver que sous l'hypothèse que f est dans L^p, alors la convergence $\phi_n \to f$ est une convergence dans L^p.

$\underline{1 \leqslant p < \infty}$ Vu que f et ϕ_n sont à valeurs positives nous avons $|f-\phi_n|^p \leqslant |f|^p$. 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

$$\lim_{n\to\infty} \int g_n = \int \lim_{n\to\infty} g_n = 0. \tag{28.130}$$

Cela revient à dire que

$$\lim_{n\to\infty} \|f-\phi_n\|_p^p = \lim_{n\to\infty} \int |f-\phi_n|^p = \int \lim_{n\to\infty} |f-\phi_n|^p = 0, \tag{28.131}$$

ce qui signifie que $\phi_n \xrightarrow{L^p(\Omega)} f$.

$\underline{\textbf{Pour } p=\infty}$ Si $f \in L^\infty(\Omega)$, alors nous pouvons prendre un représentant borné. Avec lui, nous avons $\phi_n \xrightarrow{\|.\|_\infty} f$. Avec cela nous avons, pour ϵ donné, un n assez grand pour avoir

$$N_\infty(\phi_n - f) \leqslant \|\phi_n - f\|_\infty < \epsilon. \tag{28.132}$$

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. □

28.2.8 Densité des fonctions infiniment dérivables à support compact

Définition 28.45.
Une fonction est ***étagée par rapport à*** L^p *si elle est de la forme*

$$f = \sum_{k=1}^{N} c_k \mathbb{1}_{B_k} \tag{28.133}$$

où les B_k sont des mesurables disjoints et $\mathbb{1}_{B_k} \in L^p$ pour tout k.

Lemme 28.46.
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

$$f = \sum_{k=1}^{N} c_k \mathbb{1}_{B_k} \tag{28.134}$$

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 :

$$\int_\Omega |f|^p = \int_\Omega \sum_{k=1}^{N} |c_k \mathbb{1}_{B_k}(x)|^p dx = \sum_{k=1}^{N} \int |c_k \mathbb{1}_{B_k}(x)|^p dx = \sum_{k=1}^{N} |c_k|^p \int_\Omega |\mathbb{1}_{B_k}(x)|^p dx. \tag{28.135}$$

□

Le contraire n'est pas vrai : la fonction étagée sur $\mathbb{R}$ qui vaut n sur $B(n,\frac{1}{4})$ est étagée par rapport à L^p, mais n'est pas dans L^p.

L'ensemble $C_c^\infty(\mathbb{R}^d)$ des fonctions de classe C^∞ et à support compact sur $\mathbb{R}^d$ est souvent également noté $\mathscr{D}(\mathbb{R}^d)$.

Théorème 28.47 ([411]).
Nous avons des densités emboitées. 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.

(2) Il existe une suite f_n dans $C(K,\mathbb{C})$ telle que

$$f_n \xrightarrow{L^p} \mathbb{1}_D. \tag{28.136}$$

(3) Si A est un borélien tel que $\mathbb{1}_A \in L^p(\mathbb{R}^d)$ [10] et si $\epsilon > 0$, alors il existe une suite de boréliens bornée $(D_n)_{n\in\mathbb{N}}$ tels que

$$\mathbb{1}_{D_n} \xrightarrow{L^p} \mathbb{1}_A. \tag{28.137}$$

(4) Il existe une suite φ_n dans $\mathscr{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$ telle que

$$\varphi_n \xrightarrow{L^p} \mathbb{1}_D. \tag{28.138}$$

(5) L'ensemble $\mathscr{D}(\mathbb{R}^d) = C_c^\infty(\mathbb{R}^d)$ est dense dans $L^p(\mathbb{R}^d)$ pour tout $1 \leqslant p < \infty$.

Démonstration. Nous allons montrer les choses point par point.

(1) Si $f \in L^1(\mathbb{R}^d)$, nous savons par le théorème 15.105 qu'il existe une suite f_n de fonctions étagées convergeant ponctuellement vers f telle que $|f_n| \leqslant |f|$. La proposition 28.21 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 28.46.

(2) C'est le théorème d'approximation 15.203 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

$$\|f - \mathbb{1}_D\|_p \leqslant \epsilon. \tag{28.139}$$

Ensuite, par le théorème de Weierstrass, il existe $\varphi \in C^\infty(K,\mathbb{R})$ telle que $\|f-\varphi\|_\infty \leqslant \epsilon$. Nous avons aussi

$$\|\varphi - f\|_p^p = \int_K |\varphi(x) - f(x)|^p dx \leqslant \mu(X)\|\varphi - f\|_\infty^p \leqslant \epsilon^p \mu(K). \tag{28.140}$$

Quitte à prendre un φ correspondant à un ϵ plus petit, nous avons

$$\|\varphi - f\| \leqslant \epsilon. \tag{28.141}$$

En combinant et en passant à $\epsilon/2$ nous avons trouvé une fonction $\varphi \in C^\infty(K,\mathbb{R})$ telle que

$$\|\varphi - \mathbb{1}_D\| \leqslant \epsilon. \tag{28.142}$$

(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 :

$$\int_{\mathbb{R}^d} |\mathbb{1}_{D_n}(x) - \mathbb{1}_A(x)|^p dx = \int_{\mathbb{R}^d \setminus B(0,n)} |\mathbb{1}_A(x)|^p < \epsilon, \tag{28.143}$$

c'est-à-dire que $\mathbb{1}_{D_n} \xrightarrow{L^p} \mathbb{1}_A$.

10. Je pense que cette hypothèse manque dans [411]. En tout cas je vois mal comment je pourrais justifier les différentes étapes de la preuve en prenant par exemple $A = \mathbb{R}^d$.

(5) Il suffit de remettre tout ensemble. Si $f \in L^p(\mathbb{R}^d)$, par le point (2) nous commençons par prendre σ étagée par rapport à L^p telle que

$$\|\sigma - f\|_p \leqslant \epsilon. \tag{28.144}$$

Ensuite nous écrivons σ sous la forme

$$\sigma = \sum_{k=1}^{N} c_k \mathbb{1}_{B_k} \tag{28.145}$$

et nous appliquons le point (3) à chacune des $\mathbb{1}_{B_k}$ pour trouver des boréliens bornés D_k tels que

$$\|\mathbb{1}_{D_k} - \mathbb{1}_{B_k}\|_p \leqslant \epsilon. \tag{28.146}$$

Enfin nous appliquons le point (4) pour trouver des fonctions $\varphi_k \in C_c^\infty(\mathbb{R}^d)$ telles que

$$\|\varphi_k - \mathbb{1}_{D_k}\|_p \leqslant \epsilon. \tag{28.147}$$

Il n'est pas compliqué de calculer que

$$\Big\| \sum_{k=1}^{N} c_k \varphi_k - f \Big\|_p \leqslant 2\epsilon \sum_k c_k + \epsilon. \tag{28.148}$$

□

Corollaire 28.48.
Si $1 < p < \infty$ alors la partie[11] $L^2\big([0,1]\big) \cap L^p\big([0,1]\big)$ *est dense dans* $L^p\big([0,1]\big)$.

Démonstration. Nous savons du théorème 28.47(5) que $C_c^\infty\big([0,1]\big)$ 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. □

Lemme 28.49 ([411, 391])**.**
Soit $1 \leqslant p < \infty$ et $f \in L^p(\Omega)$. Nous notons τ_v l'opérateur de translation par v :

$$\begin{aligned} \tau_v \colon L^p(\Omega) &\to L^p(\Omega) \\ f &\mapsto \big[x \mapsto f(x-v)\big]. \end{aligned} \tag{28.149}$$

Pour chaque $f \in L^p(\Omega)$, l'application

$$\begin{aligned} \tau(f) \colon \mathbb{R}^d &\to L^p(\Omega) \\ v &\mapsto \tau_v(f) \end{aligned} \tag{28.150}$$

est continue en $v = 0$, c'est-à-dire

$$\lim_{v \to 0} \|\tau_v(f) - f\|_p = 0. \tag{28.151}$$

Démonstration. Nous commençons par supposer que f est dans $\mathscr{D}(\Omega)$, et nous verrons ensuite comment généraliser.

Si $f \in \mathscr{D}(\Omega)$ 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 $\operatorname{supp}(f) \subset B(0,r)$, alors $\operatorname{supp}(f - f_i) \subset B(0, r + |v_i|)$, et l'ensemble

$$S = \overline{B\big(0, r + \max_i |v_i|\big)} \tag{28.152}$$

11. 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 topologique de L^p.

est un compact contenant les supports de tous les $f - f_i$. Le maximum existe parce que $v_i \to 0$. 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

$$M = 2 \max_{x \in \mathbb{R}^d} \{f(x)\}. \tag{28.153}$$

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 15.184) pour écrire

$$\lim_{i \to \infty} \|f - f_i\|_p^p = \lim_{i \to \infty} \int_\Omega |f(x) - f(x - v_i)|^p dx = \int_\Omega \lim_{i \to \infty} |f(x) - f(x - v_i)| dx = 0. \tag{28.154}$$

Pour $f \in L^p(\Omega)$ Soit $\epsilon > 0$, $f \in L^p(\Omega)$ et $\varphi \in \mathscr{D}(\Omega)$ tel que $\|f - \varphi\|_p \leqslant \epsilon$. Cela est possible par la densité de $\mathscr{D}(\Omega)$ dans $L^p(\Omega)$ vue en 28.47(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 τ_v étant une isométrie de L^p nous avons $\|\tau_v(\varphi) - \tau_v(f)\| = \|\varphi - f\| < \epsilon$. Nous avons tout pour majorer :

$$\|f - \tau_v(f)\| \leqslant \|f - \varphi\| + \|\varphi - \tau_v(\varphi)\| + \|\tau_v(\varphi) - \tau_v(f)\| \leqslant 3\epsilon. \tag{28.155}$$

Nous avons donc bien $\lim_{v \to 0} \|f - \tau_v(f)\| = 0$.

□

28.2.9 Approximation

Lemme 28.50 (Théorème fondamental d'approximation [243])**.**
Soit Ω un espace mesurable et $f : \Omega \to [0, \infty]$ une application mesurable. Alors il existe une suite croissante d'applications étagées $\varphi_n : \Omega \to \mathbb{R}^+$ dont la limite est f.

De plus si f est bornée, la convergence est uniforme.

Théorème 28.51 ([139])**.**
Soit I un intervalle de $\mathbb{R}$. L'espace $\mathscr{D}(I)$ 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 28.47. 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

$$\langle f, g \rangle = \int_I f \bar{g} = 0. \tag{28.156}$$

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 26.14 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

$$\int_I f_n g^+ = \int_I f_n g^-. \tag{28.157}$$

La suite étant croissante, le théorème de la convergence monotone (théorème 15.160) s'applique et nous avons

$$\lim_{n \to \infty} \int_I f_n g^+ = \int_a^b g^+, \tag{28.158}$$

de telle sorte que nous ayons, pour tout intervalle $[a,b]\subset I$ l'égalité

$$\int_a^b g^+ = \int_a^b g^-. \tag{28.159}$$

De plus ces intégrales sont finies parce que

$$\int_a^b g^+ \leqslant \int_a^b |g| = \int_I |g|\mathbb{1}_{[a,b]} = \langle |g|, \mathbb{1}_{[a,b]}\rangle \leqslant \|g\|_{L^2}\sqrt{b-a} < \infty \tag{28.160}$$

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 (φ_n) convergente vers a par le lemme 28.50. À multiples près, les fonctions φ_n sont des sommes de fonctions caractéristiques du type $\mathbb{1}_{[a,b]}$, par conséquent, en vertu de (28.159) nous avons

$$\int_I \varphi_n g^+ = \int_I \varphi_n g^-. \tag{28.161}$$

Une fois de plus nous pouvons utiliser le théorème de la convergence monotone et obtenir

$$\int_A g^+ = \int_A g^- \tag{28.162}$$

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 corollaire 15.196 du théorème de Radon Nikodym, les fonctions g^+ et g^- sont égales. □

28.3 Convolution

Définition 28.52.
Pour toutes fonctions $f,g\colon\mathbb{R}^n\to\mathbb{C}$ et pour tout $x\in\mathbb{C}$ tels que l'intégrale de droite ait un sens[12]*, nous définissons*

$$(f*g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dy. \tag{28.163}$$

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 28.53 ([263])**.**
Soient $f,g\in L^1(\mathbb{R}^n)$.

(1) Pour presque tout $x\in\mathbb{R}^n$, la fonction

$$y\mapsto g(x-y)f(y) \tag{28.164}$$

est dans $L^1(\mathbb{R}^n)$.

(2) $f*g\in L^1(\mathbb{R}^n)$.

(3) $\|f*g\|_1\leqslant\|f\|_1\|g\|_1$.

Proposition 28.54.
L'ensemble $L^1(\mathbb{R}^n)$ devient alors une algèbre de Banach.

Lemme 28.55.
Le produit de convolution est commutatif : $f*g=g*f$.

12. Attention divlgâchi : ce sera le cas pour $f,g\in L^1(\mathbb{R}^n)$ par le théorème 28.53.

Démonstration. Le théorème de Fubini (théorème 15.259) permet d'écrire

$$(f*g)(x)=\int_{\mathbb{R}^n} f(y)g(x-y)dy=\int_{-\infty}^{\infty}dy_1\dots\int_{-\infty}^{\infty}dy_n f(y)g(x-y). \tag{28.165}$$

En effectuant le changement de variable $z_i=x_i-y_i$ dans chacune des intégrales nous obtenons

$$(f*g)(x)=\int_{\mathbb{R}^n} g(z)f(x-z)dz=(g*f)(x). \tag{28.166}$$

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. □

La proposition suivante est une conséquence de l'inégalité de Minkowski sous forme intégrale de la proposition 28.37(3).

Proposition 28.56.
Si $1\leqslant p\leqslant\infty$ 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) $\|f*g\|_p\leqslant\|f\|_p\|g\|_1$.

Proposition 28.57 ([417])**.**
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 18.18) nous dit que $f*g$ est dérivable et

$$(f*g)'(x)=\frac{d}{dx}\int_{\mathbb{R}} f(t)g(x-t)dt=\int_{\mathbb{R}} f(t)g'(x-t)dt=(f*g')(x). \tag{28.167}$$

□

Corollaire 28.58.
Si $f\in L^1(\mathbb{R}^d)$ et si g est de classe C^∞, alors $f*g$ est de classe C^∞.

Démonstration. Il s'agit d'itérer la proposition 28.57. □

Lemme 28.59.
Soit $f\in L^2(I)$ telle que

$$\int_I f\varphi=0 \tag{28.168}$$

pour toute fonction $\varphi\in C_c^\infty(I)$. Alors $f=0$ presque partout sur I.

Démonstration. Nous considérons la forme linéaire

$$\begin{aligned}\phi\colon L^2(I)&\to\mathbb{C}\\ g&\mapsto\langle f,g\rangle=\int_I f\bar{g}.\end{aligned} \tag{28.169}$$

Par densité [13] nous pouvons aussi considérer une suite (φ_n) dans $C_c^\infty(I)$ convergeant dans L^2 vers f. Alors nous avons pour tout n :

$$\langle f,\varphi_n\rangle=0. \tag{28.170}$$

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. □

Ce résultat est encore valable dans les espaces L^p (proposition 28.132), mais il demande le théorème de représentation de Riesz [14].

13. Théorème 28.47(5).
14. Théorème 28.130.

28.3.1 Approximation de l'unité

Définition 28.60 ([1, 411]).
*Nous considérons $\Omega = \mathbb{R}^d$ ou $(S^1)^d$. Une **approximation de l'unité** sur Ω autour de $a \in \Omega$ est une suite (φ_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

$$\lim_{k\to\infty} \int_{\Omega\setminus V} |\varphi_k| = 0. \tag{28.171}$$

En pratique, nous allons, sur $\mathbb{R}^d$ toujours considérer des approximations de l'unité autour de 0, même si nous ne le préciserons 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 Ω est S^1, la mesure que nous considérons est $\frac{dx}{2\pi}$.

Exemple 28.61
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

$$\varphi_k(x) = k^d\varphi(kx). \tag{28.172}$$

Ici, Ω peut être $\mathbb{R}$ ou S^1. △

Le lemme suivant permet de construire des approximations de l'unité intéressantes. Nous aurons une version pour S^1 dans le lemme 28.94.

Lemme 28.62 ([411]).
Soit φ est une fonction continue et positive à support compact sur $\mathbb{R}^d$ telle que $\varphi(x) > \varphi(0)$ pour tout $x \neq 0$. Si nous posons

$$\varphi_n(x) = \left(\int \varphi(y)^n\right)^{-1} \varphi(x)^n, \tag{28.173}$$

alors la suite (φ_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 28.95.

Théorème 28.63 ([411]).
Soit (φ_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

$$(\varphi_k * g)(x_0) \to g(x_0). \tag{28.174}$$

(2) Si $g \in L^p(\mathbb{R}^d)$ $(1 \leqslant p < \infty)$ alors

$$\varphi_k * g \xrightarrow{L^p} g. \tag{28.175}$$

(3) Si g est uniformément continue et bornée, alors

$$\varphi_k * g \xrightarrow{L^\infty} g \tag{28.176}$$

Démonstration. En plusieurs points.

(1) Nous notons $d_k = (\varphi_k * g)(x_0) - g(x_0)$ et nous devons prouver que $d_k \to 0$. Vu que φ_k est d'intégrale 1 sur $\mathbb{R}^d$ nous pouvons écrire

$$d_k = \int_{\mathbb{R}^d} \varphi_k(y) g(x_0 - y) dy - \int_{\mathbb{R}^d} g(x_0)\varphi_k(y) dy, \tag{28.177}$$

et donc

$$|d_k| = \Big| \int_{\mathbb{R}^d} \big(g(x_0 - y) - g(x_0)\big)\varphi_k(y) dy \Big| \leqslant \int_{\mathbb{R}^d} \big|g(x_0 - y) - g(x_0)\big||\varphi_k(y)| dy. \tag{28.178}$$

Nous notons $M = \sup_k \|\varphi_k\|_1$, et nous considérons $\alpha > 0$ tel que

$$\big|g(x_0 - y) - g(x_0)\big| \leqslant \epsilon \tag{28.179}$$

pour tout $y \in B(0, \alpha)$. Nous nous restreignons maintenant aux k suffisamment grands pour que $\int_{\complement B(0,\alpha)} |\varphi_k(y)| dy \leqslant \epsilon$. Alors en découpant l'intégrale en $B(0, \alpha)$ et son complémentaire dans $\mathbb{R}^d$,

$$|d_k| \leqslant \epsilon M + \int_{\complement B(0,\alpha)} 2\|g\|_\infty |\varphi_k(y)| dy \leqslant \epsilon M + 2\|g\|_\infty \epsilon \leqslant \epsilon C. \tag{28.180}$$

Donc oui, nous avons $|d_k| \to 0$, et donc le premier point du théorème.

(2) Cette fois $g \in L^p(\mathbb{R}^d)$ et nous cherchons à montrer que $\|d_k\|_p \to 0$. 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 :

$$\|d_k\|_p = \Big[\int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} \big(g(x - y) - g(x)\big)\varphi_k(y) dy \Big|^p dx \Big]^{1/p} \tag{28.181a}$$

$$\leqslant \Big[\int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |g(x - y) - g(x)| \cdot |\varphi_k(y)| dy \Big)^p dx \Big]^{1/p}. \tag{28.181b}$$

À cette dernière expression nous appliquons l'inégalité de Minkowski (théorème 28.37) sous la forme (28.71) pour la mesure $d\nu(y) = |\varphi_k(y)|dy$ et $f(x, y) = g(x - y) - g(x)$:

$$\|d_k\|_p \leqslant \int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} \big|g(x - y) - g(x)\big|^p dx \Big)^{1/p} |\varphi_k(y)| dy = \int_{\mathbb{R}^d} \|\tau_y g - g\|_p |\varphi_k(y)| dy. \tag{28.182}$$

Par le lemme 28.49 nous pouvons trouver $\alpha > 0$ tel que $\|\tau_y g - g\|_p \leqslant \epsilon$ pour tout $y \in B(0, \alpha)$. Avec cela nous découpons encore le domaine d'intégration :

$$\|d_k\|_p \leqslant \int_{B(0,\alpha)} \underbrace{\|\tau_y g - g\|_p}_{\leqslant \epsilon} |\varphi_k(y)| dy + \int_{\complement B(0,\alpha)} \underbrace{\|\tau_y g - g\|_p}_{\leqslant 2\|g\|_p} |\varphi_k(y)| dy \leqslant \epsilon M + 2\epsilon \|g\|_p. \tag{28.183}$$

(3) Nous posons $d_k(x) = (\varphi_k * g)(x) - g(x)$ et nous voulons prouver que $\|d_k\|_\infty \to 0$, c'est-à-dire que $d_k(x)$ converge vers zéro uniformément en x. Nous posons aussi

$$\tau_y(g) \colon x \mapsto g(x - y). \tag{28.184}$$

En récrivant le produit de convolution, une petite majoration donne

$$|d_k(x)| \leqslant \int_{\mathbb{R}^d} \|\tau_y(g) - g\|_\infty |\varphi_k(y)| dy. \tag{28.185}$$

L'uniforme continuité de g signifie que pour tout ϵ, il existe un α tel que pour tout $y \in B(0, \alpha)$,

$$\|\tau_y(g) - g\|_\infty \leqslant \epsilon. \tag{28.186}$$

Encore une fois nous découpons le domaine d'intégration en $B = B(0, \alpha)$ et son complémentaire :

$$\|d_k\|_\infty \leqslant \int_B \underbrace{\|\tau_y(g) - g\|_\infty}_{\leqslant \epsilon} |\varphi_k(y)|dy + \int_{\complement B} \underbrace{\|\tau_y(g) - g\|_\infty}_{\leqslant 2\|g\|_\infty} |\varphi_k(y)| \tag{28.187a}$$

$$\leqslant \epsilon M + 2\|g\|_\infty \epsilon \tag{28.187b}$$

où la seconde ligne est justifiée par le choix d'un k assez grand pour que $\int_{\complement B} |\varphi_k(y)|dy \leqslant \epsilon$. Nous avons donc bien $\|d_k\|_\infty \to 0$.

□

Exemple 28.64
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 Ω n'est pas borné. En effet si $f \in L^p$, la suite donnée par

$$f_n(x) = f(x) + \frac{1}{n} \tag{28.188}$$

converge uniformément vers f, mais

$$\|f_n - f\|_p = \int_\Omega \frac{1}{n} \tag{28.189}$$

n'existe même pas si le domaine Ω n'est pas borné. △

28.3.2 Densité des polynômes trigonométriques

Définition 28.65.
*Le **système trigonométrique** donné par $\{e_n\}_{n\in\mathbb{Z}}$ est*

$$e_n(t) = \frac{1}{\sqrt{2\pi}} e^{int}. \tag{28.190}$$

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 28.66.
*Un **polynôme trigonométrique** est une fonction de la forme*

$$P(t) = \sum_{n=-N}^{N} c_n e_n(t). \tag{28.191}$$

Définition 28.67 (Coefficients de Fourier)**.**
Pour toute fonction pour laquelle ça a un sens (que ce soit des fonctions L^2 ou non), nous posons

$$c_n(f) = \langle f, e_n \rangle. \tag{28.192}$$

*Ces nombres sont les **coefficients de Fourier** de f.*

Ces trois définitions n'ont a priori aucun rapport entre elles, et rien en particulier ne devrait vous faire penser à une égalité du type

$$f(x) = \sum_{n=-\infty}^{\infty} c_n(f) e_n(x). \tag{28.193}$$

Nous avons toutefois quelque liens.

Lemme 28.68.
Deux petits résultats simples mais utiles à propos des polynômes trigonométriques.

(1) Si $f \in L^1(S^1)$, alors nous avons la formule

$$f * e_n = c_n(f)e_n. \tag{28.194}$$

(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 :

$$(f * e_n)(x) = \int_0^{2\pi} f(x-t)e_n(t) \tag{28.195a}$$

En ce qui concerne le second point, nous notons $P = \sum_{k=-N}^{N} P_k e_k$, et par linéarité de la convolution,

$$f * P = \sum_{k=-N}^{N} P_k f * e_k = \sum_{k=-N}^{n} P_k c_k(f) e_k, \tag{28.196}$$

qui est encore un polynôme trigonométrique. □

Exemple 28.69
Sur S^1 nous construisons alors l'approximation de l'unité basée sur la fonction $1 + \cos(x)$ et le lemme 28.62. Cette fonction est évidemment un polynôme trigonométrique parce que

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2}. \tag{28.197}$$

Ensuite les puissances le sont aussi à cause de la formule du binôme :

$$\big(1 + \cos(x)\big)^n = \sum_{k=0}^{n} \binom{n}{k} \cos^n(x), \tag{28.198}$$

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 (φ_n) formée de polynômes trigonométrique.

Ce qui fait la spécificité des polynômes trigonométriques est qu'ils sont à la fois stables par convolution (lemme 28.68) 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. △

Théorème 28.70.
Les polynôme trigonométriques sont dense dans $L^p(S^1)$ pour $1 \leqslant p < \infty$.

Démonstration.

$$\varphi_k * f \xrightarrow{L^p} f \tag{28.199}$$

par le théorème 28.63. Nous avons donc convergence L^p d'une suite de polynômes trigonométrique, ce qui prouve que l'espace de polynômes trigonométriques est dense dans $L^p(S^1)$. □

Remarque 28.71.
Deux remarques.

— Il n'est pas possible que les polynômes trigonométriques soient dense dans L^∞ parce qu'une limite uniforme de fonctions continues est continue (c'est le théorème 13.280). Donc les polynômes trigonométriques ne peuvent engendrer que des fonctions continues.

— Nous donnerons au théorème 29.6 une démonstration indépendante de la densité des polynômes trigonométriques dans $L^p(S^1)$.

28.4 Espaces L^2, généralités

L'espace L^2 est l'espace L^p définit en 28.8 avec $p = 2$. Cependant il possède une propriété extraordinaire par rapport aux autres L^p, c'est que la norme $|.|_2$ dérive d'un produit scalaire. Il sera donc un espace de Hilbert.

28.72.
Nous en rappelons la construction. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Nous considérons l'opération

$$\langle f, g\rangle = \int_\Omega f(\omega)\overline{g(\omega)}d\mu(\omega) \tag{28.200}$$

et la norme associée

$$\|f\|_2 = \sqrt{\langle f, f\rangle}. \tag{28.201}$$

Nous considérons l'ensemble

$$\mathcal{L}^2(\Omega, \mu) = \{f : \Omega \to \mathbb{C} \text{ tel que } \|f\|_2 < \infty\} \tag{28.202}$$

et la relation d'équivalence $f \sim g$ si et seulement si $f(x) = g(x)$ pour μ-presque tout x.

Et enfin, nous considérons le quotient

$$L^2(\Omega, \mu) = \mathcal{L}^2(\Omega, \mu)/\sim . \tag{28.203}$$

Lemme 28.73.
Soit un espace mesuré[15] $(\Omega, \mathcal{A}, \mu)$.

(1) Pour tout $f, g \in L^2(\Omega, \mathcal{A}, \mu)$, le produit

$$\langle f, g\rangle = \int_\Omega f\bar{g}\, d\mu \tag{28.204}$$

est bien défini et est un nombre complexe[16].

(2) L'opération $(f, g) \mapsto \langle f, g\rangle$ est un produit hermitien[17].

(3) Le couple $\big(L^2(\Omega, \mathcal{A}, \mu), \langle ., .\rangle\big)$ est un espace de Hilbert[18].

Démonstration. Que $L^2(\Omega)$ soit un espace vectoriel est un cas particulier de la proposition 28.17. Voyons cette histoire de produit scalaire.

Pour de vraies fonctions Nous commençons par analyser l'intégrale (28.204) 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 28.33) 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.

Passage aux classes Ensuite nous montrons que la formule passe au quotient. Pour cela, nous considérons des fonctions α et β nulles presque partout et nous regardons le produit de $f_1 = f + \alpha$ par $g_1 = g + \beta$:

$$\langle f_1, g_1\rangle = \int fg + \beta f + \alpha g + \alpha\beta. \tag{28.205}$$

Les fonctions βf, αg et $\alpha\beta$ 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.

15. Est-ce qu'il ne faudrait pas un peu plus d'hypothèses, comme σ-fini par exemple ? Vérifiez et écrivez-moi quand vous avez la réponse.
16. Par opposition au fait que ce serait l'infini.
17. Définition 11.8. 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.
18. Définition 26.1.

Produit hermitien Pour vérifier que la formule est un produit hermitien, le seul point non évidement est de prouver que $\langle f, f\rangle = 0$ implique $f = 0$. Cela découle du fait que

$$\langle f, f\rangle = \int_\Omega |f|^2. \tag{28.206}$$

La fonction $x \mapsto |f(x)|^2$ vérifie les hypothèses du lemme 15.179. Par conséquent $|f(x)|^2$ est presque partout nulle.

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 28.40) et dont la norme dérive d'un produit scalaire ou hermitien. Nous sommes donc bien dans la définition 26.1.

□

28.74.
Ces espaces seront utilisés pour de nombreuses applications. Nous en aurons besoin pour plusieurs combinaisons d'ensembles Ω et de mesures μ.

- Pour $\mathbb{R}^d$
- Pour S^1
- Pour $[a, b]$
- Pour $[0, 2\pi[$
- Pour $[-T, T[$

Le premier est non compact et il est raisonnable de penser qu'il sera foncièrement différents 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 28.73 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 $\mathrm{Fun}(\Omega, \mathbb{C})$, l'ensemble de toutes les fonctions sur Ω à valeurs dans $\mathbb{C}$. Nous verrons plus tard les espaces de fonctions sur lesquels tout a un sens.

28.5 L'espace $L^2(\mathbb{R}^d)$

La mesure est celle de Lebesgue. Le produit de convolution est donné, pour $f, g \in \mathrm{Fun}(\mathbb{R}^d, \mathbb{C})$, par

$$(f * g)(x) = \int_{\mathbb{R}^d} f(y)g(x - y)dy \tag{28.207}$$

Certaines de ses propriétés ont déjà été vues dans le théorème 28.53.

En ce qui concerne le système trigonométrique, pour tout $\xi \in \mathbb{R}^d$ nous définirions bien

$$e_\xi(x) = e^{i\xi \cdot x}, \tag{28.208}$$

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-compacité de $\mathbb{R}^d$ allait un peu changer les choses par rapport aux autres, je ne rigolais pas.

28.6 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,

$$S^1 = \{e^{it}\}_{t\in\mathbb{R}}, \tag{28.209}$$

sans garanties que cette paramétrisation 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 21.9. 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 15.255(5).

28.6.1 Espace mesuré

Plusieurs choses sont déjà faites.

— Les boréliens de S^1 sont décrits dans la proposition 19.64,

— la tribu de Lebesgue de S^1 est décrite dans la proposition 19.66. Non, ce n'est pas la tribu induite de la tribu de Lebesgue de $\mathbb{C}$.

Proposition 28.75 (Espaces de fonctions sur S^1[1])**.**
Soit l'espace mesuré $\big(S^1, \mathcal{L}eb(S^1), \mu\big)$.

(1) La formule

$$\langle f,g\rangle = \int_{S^1} f\bar{g}d\mu \tag{28.210}$$

est un produit hermitien[19] *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

$$\begin{aligned}\phi\colon L^2(S^1) &\to L^2\big([0,2\pi[\big)\\ f &\mapsto \frac{1}{\sqrt{2\pi}} f\circ\varphi.\end{aligned} \tag{28.211}$$

est une bijection isométrique (isomorphisme d'espaces de Hilbert)

$$L^2(S^1,\mathcal{L}eb(S^1),\mu) = L^2\big(]0,2\pi[,\mathcal{L}eb(\mathbb{R}),\lambda\big) \tag{28.212}$$

où nous avons fait un minuscule abus de notations : ici $\mathcal{L}eb(\mathbb{R})$ *est en réalité la tribu induite sur* $]0,2\pi[$.

19. Définition 11.8.

Démonstration. Le fait que la formule (28.210) donne bien un produit hermitien est le lemme 28.73. 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 [20]

$$\begin{aligned}\phi\colon L^2(S^1) &\to L^2\big([0,2\pi[\big)\\ f &\mapsto \frac{1}{\sqrt{2\pi}} f\circ\varphi.\end{aligned} \tag{28.213}$$

Injection Si $\phi(f)=\phi(g)$, alors pour tout $x\in[0,2\pi[$ nous avons $f\big(\varphi(x)\big)=g\big(\varphi(x)\big)$. Vu que $\varphi\colon[0,2\pi[\to S^1$ est une bijection nous avons alors $f(s)=g(s)$ pour tout $s\in S^1$.

Surjection Si $f\in L^2\big(]0,2\pi]\big)$, nous posons $g\colon S^1\to\mathbb{C}$ par

$$g(s)=\sqrt{2\pi}f\big(\varphi^{-1}(s)\big). \tag{28.214}$$

Nous avons alors bien $\phi(g)(x)=f(x)$.

Isométrie Nous montrons que ϕ préserve le produit scalaire :

$$\begin{aligned}\big\langle\phi(f),\phi(g)\big\rangle &= \int_0^{2\pi}\phi(f)(x)\overline{\phi(g)(x)}d\lambda(x) &\text{(28.215a)}\\ &= \frac{1}{2\pi}\int_0^{2\pi}(f\circ\varphi)(x)\overline{(g\circ\varphi)(x)}\,d\lambda(x) &\text{(28.215b)}\\ &= \frac{1}{2\pi}\int_0^{2\pi}(fg)\circ\varphi\,d\lambda &\text{(28.215c)}\end{aligned}$$

Pour la suite nous devons invoquer la proposition 15.254 pour passer d'une intégrale sur $\big([0,2\pi[,\mathcal{L}\text{eb}\big([0,2\pi[\big),\lambda\big)$ à une intégrale sur $\big(S^1,\mathcal{L}\text{eb}(S^1),\mu\big)$. La première condition de cette proposition est que $\mathcal{L}\text{eb}(S^1)=\varphi\big(\mathcal{L}\text{eb}([0,2\pi[)\big)$. Cela est la proposition 19.66(2). La condition sur la mesure dans la proposition 15.254 n'est vraie ici qu'à un facteur 2π près. Nous avons :

$$\int_{[0,2\pi[} f d\lambda = 2\pi\int_{S^1}(f\circ\varphi^{-1})d\mu. \tag{28.216}$$

Nous continuons le calcul (28.215) :

$$\big\langle\phi(f),\phi(g)\big\rangle=\frac{1}{2\pi}\int_0^{2\pi}(fg)\circ\varphi\,d\lambda=\int_{S^1}f\bar{g}d\mu=\langle f,g\rangle. \tag{28.217}$$

□

28.6.2 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 7.46, l'espace S^1 muni de sa topologie est un espace topologique compact.

Nous pouvons donc sans crainte affirmer que toute fonction continue $f\colon S^1\to\mathbb{K}$ est bornée et atteint ses bornes.

Proposition-définition 28.76.
Soit la formule

$$d(e^{ix},e^{iy})=\inf_{k\in\mathbb{Z}}|x-y+2k\pi|. \tag{28.218}$$

(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).

20. Notez que cette définition passe aux classes. Nous le répéterons pas.

(3) La formule définit une distance[21] *sur* S^1.

Nous considérons sur S^1 *la topologie* τ_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 $l, l' \in \mathbb{Z}$ (corollaire 19.33). Nous avons alors $|x' - y' + 2k\pi| = |x - y + 2\pi(k + l - l')|$ et

$$\inf_{k\in\mathbb{Z}} |x' - y' + 2k\pi| = \inf_{k\in\mathbb{Z}} |x - y + 2k\pi|. \tag{28.219}$$

(2) Quels que soient x et y fixés, nous avons

$$\lim_{k\to\pm\infty} |x - y + 2k\pi| = \infty. \tag{28.220}$$

Donc l'infimum est forcément atteint par un $k \in \mathbb{Z}$.

(3) Pour la distance, il y a plusieurs points à prouver.

— Pour tout $z, z' \in S^1$ nous avons $d(z, z') \geqslant 0$ 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

$$|x - y + 2k\pi| = |y - x - 2k\pi| = |y - x + 2k'\pi| \tag{28.221}$$

en posant $k' = -k$. L'infimum étant pris sur $k \in \mathbb{Z}$, nous avons al symétrique $d(e^{ix}, e^{iy}) = d(e^{iy}, e^{ix})$.

— 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

$$\begin{aligned}
d(z_1, z_3) &= |x_1 - x_3 + 2k_1\pi| = |x_1 - x_2 + x_2 - x_3 + 2k_1\pi| && \text{(28.222a)}\\
&= \inf_{k\in\mathbb{Z}} |(x_1 - x_2 + 2k_2\pi) + (x_2 - x_3 + 2k_3\pi) + 2k\pi| && \text{(28.222b)}\\
&= \inf_{k\in\mathbb{Z}} \Big(|x_1 - x_2 + 2k_1\pi| + |x_2 - x_3 + 2k_3\pi| + 2k\pi\Big) && \text{(28.222c)}\\
&= d(z_1, z_3) + d(z_2, z_3) && \text{(28.222d)}
\end{aligned}$$

parce que le dernier infimum est réalisé par $k = 0$.

□

Le cercle est bien connu pour être symétrique et en particulier avoir une symétrie sous les rotations. Nous allons voir quelque résultats qui vont dans le sens de dire que la distance définie sur S^1 respecte cette symétrie.

Lemme 28.77.
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*

$$d(e^{is}a, e^{is}b) = d(a, b). \tag{28.223}$$

(2) Les boules sont préservées sous les rotations[22]*, c'est-à-dire que*

$$e^{is}B_d(a, r) = B_d(e^{is}a, r). \tag{28.224}$$

21. Définition 7.87.
22. Pas chaque boule séparément, mais l'ensemble des boules

(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

$$d(a,b) = \inf_{k\in\mathbb{Z}} |x - y + 2k\pi| = \inf_{k\in\mathbb{Z}} |(x-s) - (y-s) + 2k\pi| = d(e^{is}a, e^{is}b). \tag{28.225}$$

Et de un.

(2) Il faut une inclusion dans chaque sens.

$e^{is}B_d(a,r) \subset B_d(e^{is}a,r)$ 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,

$$d(b,e^{is}) = d(e^{-is}b,a) = d(b',a) < r. \tag{28.226}$$

Donc $b \in B_d(e^{is}a,r)$.

$B_d(e^{is}a,r) \subset e^{is}B_d(a,r)$ 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,

$$d(e^{-is}b,a) = d(b,e^{is}a) < r. \tag{28.227}$$

Donc oui, $e^{-is}b \in B(a,r)$.

(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

$$e^{is}aB_d(a',r) \subset e^{is}A, \tag{28.228}$$

et comme $e^{is}B_d(a',r) = B_d(e^{is}a',r) = B_d(a,r)$ nous avons bien

$$B_d(a,r) \subset e^{is}A. \tag{28.229}$$

□

Nous allons voir maintenant quelque 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 28.1, 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.

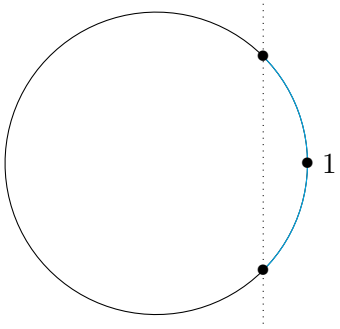


FIGURE 28.1 – Un voisinage de 1 dans S^1.

Ce lemme-ci montre que $B_d(1,r)$ est une partie de S^1 qui s'étale symétriquement autour de 1.

Lemme 28.78.
Soit l'application

$$\begin{aligned}\varphi\colon \mathbb{R} &\to S^1\\ x &\mapsto e^{ix}.\end{aligned} \tag{28.230}$$

Nous avons $B_d(1,r) = \varphi\big(]-r,r[\big)$.

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 $b \in \varphi\big(]-r,r[\big)$.

Dans l'autre sens, si $y \in]-r,r[$, alors

$$d(1,e^{iy}) = \inf_{k\in\mathbb{Z}} |y + 2k\pi| \leqslant |y| < r. \tag{28.231}$$

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 (28.231) 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 28.1.

Lemme 28.79.
Soit $r \in [0,\pi]$. *Nous avons*

$$B_d(1,r) = S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}. \tag{28.232}$$

Démonstration. Si $r = \pi$, alors $B(1,r) = S^1\backslash\{-1\}$, alors que $\cos(\pi) = -1$.

Si $r < \pi$, alors nous partons de la formule (19.70) qui dit que $e^{ir} = \cos(e) + i\sin(r)$. D'après le lemme 28.78, 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, donc si $y < 0$ alors

$$\cos(y) > \cos(-r) = \cos(r). \tag{28.233}$$

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 $B_d(1,r) \subset S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}$.

Lançons nous dans la preuve de l'inclusion inverse.

Soit $x > \cos(r)$. Si $x + iy \in S^1$, nous avons $x + iy = e^{is} = \cos(s) + i\sin(s)$ pour un certain $s \in [-\pi,\pi[$. Notons que $s = -\pi$ correspondrait au point $-1 \in S^1$, qui est exclu de notre étude parce que nous supposons $r < \pi$. Donc $s \in]-\pi,\pi[$.

Nous avons donc $\cos(r) < x = \cos(s)$. Et voila. □

Proposition 28.80.
La topologie τ_d *sur* S^1 [23] *est la topologique induite depuis* $\mathbb{C}$.

Démonstration. Nous notons τ_i la topologie induite (c'est-à-dire l'ensemble des ouverts) et τ_d la topologie de la distance fraichement 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.

$\tau_i \subset \tau_d$ Un élément général de τ_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 τ_d contenant a et contenu dans $\mathcal{O} \cap S^1$; cela prouvera que $\mathcal{O} \cap S^1$ est ouvert de τ_d par le théorème 7.4.

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

$$d(a,b) = \inf_{k\in\mathbb{Z}} |x - y + 2k\pi| \leqslant r. \tag{28.234}$$

23. Définition 28.76.

Quitte à redéfinir x ou y nous pouvons supposer que l'infimum est atteint en $k = 0$. En utilisant la proposition 22.15 nous majorons :

$$|a - b| = |e^{ix} - e^{iy}| \leqslant |x - y| = d(a, b) \leqslant r. \tag{28.235}$$

Donc nous avons bien $b \in B_{\mathbb{C}}(a, r)$ dès que $b \in B_d(a, r)$.

$\tau_d \subset \tau_i$ Ce sens est plus délicat parce que, si nous voulons suivre les mêmes pas que le premier sens, nous devrons nous appuyer sur la continuité de l'application $\ln\colon \mathbb{C}^* \to \mathbb{C}$, laquelle n'est pas vraie en -1 (voir par exemple le lemme 27.58).

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.

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 28.79 donne

$$B_d(1, r) = S^1 \cap \{x + iy \text{ tel que } x > \cos(r)\}. \tag{28.236}$$

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

$$1 - \delta > 1 - (1 - \cos(r)) = \cos(r), \tag{28.237}$$

ce qui prouve que la partie de $\mathcal{O}$ qui est dans S^1 est bien dans $B_d(1, r)$.

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 $e^{-ix}B_d(d, r) \subset e^{-ix}A$ et donc, en tenant compte du lemme 28.77(2) :

$$B_d(1, r) \subset e^{-ix}A. \tag{28.238}$$

Par le premier point, il existe un ouvert $\mathcal{O}$ de $\mathbb{C}$ tel que $1 \in \mathcal{O}$ et

$$\mathcal{O} \cap S^1 \subset B_d(1, r) \subset e^{-ix}A. \tag{28.239}$$

Nous avons évidemment que $a \in e^{ix}\mathcal{O}$ et

$$e^{ix}(\mathcal{O} \cap S^1) \subset e^{ix}B_d(1, r) \subset A. \tag{28.240}$$

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 τ_i.

□

Lemme 28.81 ([1]).
Deux résultats de limites dans S^1.

(1) Pour tout $a_0 \in S^1$, nous avons

$$\lim_{s \to 1} d(a, as) = 0. \tag{28.241}$$

(2) Si $f\colon S^1 \to \mathbb{C}$ est continue en $a \in S^1$, alors

$$\lim_{s \to 1} f(as) = f(a). \tag{28.242}$$

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

$$d(a, as) = \inf_{k \in \mathbb{Z}} |x - (x + y) + 2k\pi| = \int_{k \in \mathbb{Z}} |y + 2k\pi|. \tag{28.243}$$

Voila 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 28.78. Pour un tel s nous avons

$$d(a, sa) = \inf_{k\in\mathbb{Z}} |y + 2k\pi| \leqslant |y| < \delta < \epsilon. \tag{28.244}$$

Nous nous avons trouvé $\delta > 0$ tel que $s \in B(1, \delta)$ implique $d(a, as) < \epsilon$. Cela est la limite que nous devions 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 δ est le point (1) de ce lemme.

Avec tout cela nous avons $|g(as)-g(a)| < \epsilon$ dès que $s \in B(1, \delta)$. Cela est la limite $\lim_{s\to 1} g(as) = g(a)$ que nous voulions prouver.

□

Lemme 28.82.
Pour $s \in S^1$, nous considérons l'application

$$\begin{aligned} \alpha_s\colon \operatorname{Fun}(S^1) &\to \operatorname{Fun}(S^1) \\ \alpha_s(g)(u) &= g(u\bar{s}) - g(u). \end{aligned} \tag{28.245}$$

Quelque propriétés avec $1 \leqslant p < \infty$:

(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

$$\lim_{s\to 1} \alpha_s(f) = 0 \tag{28.246}$$

dans $L^p(S^1)$.

Démonstration. D'abord un calcul de norme :

$$\|\alpha_s(f)\|_p^p = \int_{S^1} |f(u\bar{s}) - f(u)|^p du \leqslant \int_{S^1} |f(u\bar{s})|^p du + \int_{S^1} |f(u)|^p du = 2\|f\|_p^p. \tag{28.247}$$

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 $\|f\|_\infty$ nous avons $|\alpha_s(f)| \leqslant 2M$. Donc la fonction constante

$$\begin{aligned} g\colon S^1 &\to \mathbb{C} \\ u &\mapsto 2M \end{aligned} \tag{28.248}$$

est une fonction intégrable sur S^1 qui majore $|\alpha_s(f)|$ uniformément en s. Soit une suite $s_i \to 1$ dans S^1, et posons $f_i = \alpha_{s_i}(f)$. Alors nous avons

$$\|f_i\|_p^p = \int_{S^1} |f_i(u)|^p du \tag{28.249}$$

et aussi $|f_i|^p \leqslant (2M)^p$. Le théorème de la convergence dominée de Lebesgue 15.184 nous permet de permuter limite et intégrale :

$$\lim_{i\to\infty} \|f_i\|_p^p = \int_{S^1} \lim_{i\to\infty} |f_i(u)|^p du. \tag{28.250}$$

Mais

$$\lim_{i\to\infty} f_i(u) = \lim_{i\to\infty} \alpha_{s_i}(f)(u) = \lim_{i\to\infty} \big(f(u\bar{s}_i) - f(u)\big) = 0. \tag{28.251}$$

La dernière limite est due au fait que $\lim_{s\to 1} g(us) = g(u)$ (lemme 28.81(2)). □

28.6.3 Système trigonométrique

Définition 28.83.
*La **famille trigonométrique** sur S^1 est l'ensemble de fonctions $\{e_n\}_{n\in\mathbb{Z}}$ données par*

$$\begin{aligned} e_n\colon S^1 &\to \mathbb{C} \\ z &\mapsto z^n \end{aligned} \tag{28.252}$$

*avec $n \in \mathbb{Z}$. Un **polynôme trigonométrique** est une application $S^1 \to \mathbb{C}$ de la forme*

$$\sum_{k=-n}^{n} a_k e_k \tag{28.253}$$

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 $[0, 2\pi[$.

Proposition 28.84.
La famille trigonométrique est une famille orthonormale pour le produit scalaire $L^2(S^1, \mathcal{A}, \mu)$.

Démonstration. En utilisant la proposition 28.75(3) nous avons :

$$\langle e_n, e_n\rangle = \int_{S^1} e_n\overline{e_n} = \frac{1}{2\pi}\int_{[0,2\pi[} e_n\big(\varphi(x)\big)\overline{e_n\big(\varphi(x)\big)} \tag{28.254a}$$

$$= \frac{1}{2\pi}\int_{[0,2\pi[} e^{inx}e^{-inx}dx = \frac{1}{2\pi}\int_0^{2\pi} 1dx = 1. \tag{28.254b}$$

Et nous avons également, pour $m \neq n$:

$$\langle e_n, e_m\rangle = \frac{1}{2\pi}\int_{[0,2\pi[} e^{i(n-m)x}dx = \frac{1}{2\pi}\left[\frac{1}{i(n-m)e^{i(n-m)x}}\right]_0^{2\pi} = 0. \tag{28.255}$$

□

Remarque 28.85.
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 19.181.

Remarque 28.86.
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 $n \in \mathbb{N}$.

Proposition 28.87.
Les polynômes trigonométriques forment une partie dense dans $\big(C(S^1, \mathbb{C}), \|.\|_\infty\big)$.

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 13.303.

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

$$P = \sum_{k=-n}^{n} a_k e_k, \tag{28.256}$$

alors $\bar{P} = \sum_{k=-n}^{n} \overline{a_k e_k a} = \sum_{k=-n}^{n} \overline{a_k} e_{-k}$ qui est encore un polynôme trigonométrique. □

Définition 28.88.
Si nous avons une fonction $f\colon S^1 \to \mathbb{C}$, nous définissons ses ***coefficients de Fourier*** *par*

$$c_n(f) = \langle f, e_n \rangle \tag{28.257}$$

pourvu que l'intégrale existe.

28.6.4 Convolution

La convolution sur $\mathbb{R}^n$ est donnée par la définition 28.52. Nous voyons maintenant comment cela s'adapte à S^1.

Définition 28.89.
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*

$$(f * g)(z) = \int_{S^1} f(s)g(z\bar{s})d\mu(s). \tag{28.258}$$

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 (28.258) 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 28.90 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 (28.258), la variable s est vraiment une variable muette. Cette formule aurait également pu être écrite

$$(f * g)(z) = \int_{S^1} \big[s \mapsto f(s)g(z\bar{s})\big]d\mu. \tag{28.259}$$

Lemme 28.90 ([418]).
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

$$\begin{aligned} \psi\colon S^1 \times S^1 &\to \mathbb{C} \\ (z, s) &\mapsto f(s)g(z\bar{s}). \end{aligned} \tag{28.260}$$

$\psi \in L^1(S^1 \times S^1)$ Nous utilisons le corollaire 15.258, et pour cela nous calculons les intégrales en chaîne[24] :

$$\int_{S^1} \left[\int_{S^1} |f(s)g(z\bar{s})|dz\right]ds = \int_{S^1} |f(s)| \underbrace{\left[\int_{S^1} |g(z\bar{s})|dz\right]}_{=A<\infty} ds \tag{28.261a}$$

$$= A \int_{S^1} |f(s)ds| \tag{28.261b}$$

$$< \infty. \tag{28.261c}$$

Le fait que $A < \infty$ provient directement de l'hypothèse $g \in L^1(S^1)$[25].

Par le corollaire sus-cité nous avons bien $\psi \in L^1(S^1 \times S^1)$.

24. 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.

25. Avec un changement de variables $z \mapsto z\bar{s}$ que je vous conseille d'être capable de justifier.

Et par Fubini Le théorème de Fubini 15.259(1) nous renseigne que pour presque tout $z \in S^1$, l'application

$$s \mapsto \psi(z, s) \tag{28.262}$$

est dans $L^1(S^1)$. Et la partie 15.259(2) ajoute que l'application

$$z \mapsto \int_{S^1} \psi(s, z) ds \tag{28.263}$$

est également $L^1(S^1)$.

Conclusion L'application donnée en (28.263) est précisément $(f * g)$. Donc $f * g \in L^1(S^1)$.

□

Lemme 28.91 ([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 28.90 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 Heine13.89. 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 28.77(2). Dans ce cas nous avons aussi

$$|g(\bar{s}z_0) - g(\bar{s}z)| < \epsilon. \tag{28.264}$$

Un peu de calcul maintenant. D'une part

$$(f * g)(z_0) - (f * g)(z) = \int_{S^1} f(s)\big(g(z_0\bar{s}) - g(z\bar{s})\big) ds, \tag{28.265}$$

et donc

$$|(f * g)(z_0) - (f * g)(z)| \leqslant \int_{S^1} |f(s)| \underbrace{\big|g(z_0\bar{s}) - g(z\bar{s})\big|}_{<\epsilon} ds \tag{28.266a}$$

$$\leqslant \epsilon \int_{S^1} |f| \tag{28.266b}$$

$$= A\epsilon \tag{28.266c}$$

pour une certaine constante A ne dépendant pas de z_0.

Nous avons prouvé que pour tout $\epsilon > 0$, il existe un δ tel que $z \in B(z_0, \alpha)$ implique

$$|(f * g)(z_0) - (f * g)(z)| \leqslant A\epsilon, \tag{28.267}$$

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 28.92.
Si $f \in L^1(S^1)$, nous avons

$$f * e_n = c_n(f) e_n. \tag{28.268}$$

Démonstration. Il s'agit d'un bon calcul. En considérant $z = e^{i\theta}$ nous avons

$$(f * e_n)(z) = \int_{S^1} f(s)e_n(z\bar{s})ds \tag{28.269a}$$
$$= \frac{1}{2\pi}\int_{[0,2\pi[} f(e^{ix})e_n(e^{i(\theta-x)})dx \tag{28.269b}$$
$$= e^{in\theta}\frac{1}{2\pi}\int_{[0,2\pi[} f(e^{ix})e^{-inx}dx \tag{28.269c}$$
$$= e_n(e^{i\theta})\frac{1}{2\pi}\int_{[0,2\pi[} f(e^{ix})\overline{e_n(e^{ix})}dx \tag{28.269d}$$
$$= e_n(z)\int_{S^1} f(s)\overline{e_n(s)}ds \tag{28.269e}$$
$$= e_n(z)\langle f, e_n\rangle. \tag{28.269f}$$

Donc $(f * e_n)(z) = \langle f, e_n\rangle e_n(z)$, c'est-à-dire que

$$f * e_n = c_n(f)e_n. \tag{28.270}$$

□

Lemme 28.93.
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,

$$f * P = \sum_{k=-n}^{n} a_k f * e_k = \sum_k a_k c_k(f)e_k \tag{28.271}$$

où nous avons également utilisé la proposition 28.92. Nous avons donc un polynôme trigonométrique dont les coefficients sont $a_k c_k(f)$ au lieu de a_k. □

28.6.5 Approximation de l'unité

Lemme 28.94 ([411])**.**
Soient une fonction continue $f\colon S^1 \to [0,\infty[$ et $a \in S^1$ telle que $f(z) < f(a)$ pour tout $z \in S^1\backslash\{a\}$. Alors la suite de fonctions $f_n\colon S^1 \to \mathbb{R}$ donnée par

$$f_n(z) = \left(\int_{S^1} f^n\right)^{-1} f(z)^n \tag{28.272}$$

est une approximation de l'unité[26] *autour de a.*

Démonstration. En plusieurs points, dont d'abord une série de vérifications pour voir que la formule a un sens.

Strictement positive 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.

f^n est intégrable sur S^1 La fonction f^n est dans les hypothèses de la proposition 15.191 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).

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

$$\int_K f^n \geqslant \int_A f^n > 0. \tag{28.273}$$

Cela pour dire que l'inverse dans (28.272) ne pose pas de problèmes.

26. Définition 28.60.

Norme Vu que toutes les fonctions tant f que f_n sont positives, les valeurs absolues ne jouent aucun rôle et nous avons

$$\|f_n\|_1 = \int_{S^1} f_n d\mu = \left(\int_{S^1} f^n\right)^{-1} \int_{S^1} |f(z)|^n d\mu = 1. \tag{28.274}$$

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

$$A_k = \{x \in S^1 \backslash V \text{ tel que } f(x) \geqslant t_k\}. \tag{28.275}$$

Cet ensemble est contenu dans S^1 et est donc borné (pour la métrique de S^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 7.4 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 $f < t_k$.

A_k est compact L'espace S^1 est compact, par exemple grâce au lemme 7.46. La partie A_k est fermée dans le compact S^1, donc elle est compacte par le lemme 7.59.

$A_{k+1} \subset A_k$ Si $x \in A_{k+1}$, alors $f(x) \geqslant t_{k+1} > t_k$. Donc $f(x) > t_k$ et $x \in A_k$.

Intersection vide Si $x \in \bigcap_{k\in\mathbb{N}} A_k$, alors $f(x) \geqslant t_k$ pour tout k. En passant à la limite et en sachant que $\lim_{k\to\infty} t_k = f(a)$, nous avons $f(x) \geqslant f(a)$. Par hypothèse, cela n'est pas. Donc

$$\bigcap_{k\in\mathbb{N}} A_k = \varnothing. \tag{28.276}$$

Nous avons, dans un compact, des fermés emboîtés dont l'intersection est vide. Le corollaire 7.53 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

$$\{x \in S^1 \backslash V \text{ tel que } f(x) \geqslant \delta\} = \varnothing, \tag{28.277}$$

c'est-à-dire que sur $S^1\backslash V$, nous avons $f < \delta$ et donc

$$\int_{S^1\backslash V} f(s)^n < \int_{S^1\backslash V} \delta^n = \text{Vol}(S^1\backslash V)\delta^n \tag{28.278}$$

où $\text{Vol}(S^1\backslash V) = \int_{S^1\backslash V} 1 = \mu'(S^1\backslash V)$ est une constante réelle strictement positive.

Nous avons aussi $\delta < f(a)$ parce que δ est un des t_k (et que cette suite croissante converge vers $f(a)$ sans l'atteindre par hypothèse). Soit δ_1 tel que $\delta < \delta_1 < f(a)$ [27].

Nous posons

$$W = \{x \in S^1 \text{ tel que } f(x) > \delta_1\}. \tag{28.279}$$

Cet ensemble n'est pas vide parce qu'il contient a et est ouvert parce que f est continue. Nous avons

$$\int_{S^1} f(s)^n ds \geqslant \int_W f(s)^n ds \geqslant \delta_1^n \text{ Vol}(W). \tag{28.280}$$

27. Dans [411], il prend $\delta < \delta_1 < 1$ et je crois qu'il aurait dû écrire $\varphi(0)$ au lieu de 1.

Nous avons donc déjà ces deux inégalités :

$$\int_{S^1\backslash V} f(s)^n <= \mathrm{Vol}(S^1\backslash V)\delta^n \tag{28.281}$$

et

$$\int_{S^1} f(s)^n ds \geqslant \delta_1^n \mathrm{Vol}(W). \tag{28.282}$$

En ce qui concerne les fonctions f_n que nous voulions étudier,

$$f_n(z) = \left(\int_{S^1} f(s)^n ds\right)^{-1} f(z)^n \leqslant \big(\mathrm{Vol}(W)\delta_1^n\big)^{-1} f(z)^n, \tag{28.283}$$

et donc

$$\int_{S^1\backslash V} f_n \leqslant \big(\mathrm{Vol}(W)\delta_1^n\big)^{-1}\mathrm{Vol}(S^1\backslash V)\sigma^n = \frac{\mathrm{Vol}(S^1\backslash V)}{\mathrm{Vol}(W)}\left(\frac{\delta}{\delta_1}\right)^n. \tag{28.284}$$

Étant donné que $\delta < \delta_1$, nous avons $(\delta/\delta_1)^n \to 0$. Donc aussi

$$\lim_{n\to\infty}\int_{S^1\backslash V} f_n(z)dz = 0. \tag{28.285}$$

□

Le théorème suivant est une version pour S^1 du théorème 28.63. Le produit de convolution dans S^1 est la définition 28.89.

Théorème 28.95 ([411, 1])**.**
Soient (φ_k) une approximation de l'unité sur $\Omega = S^1$ ainsi qu'une fonction $g\colon S^1 \to \mathbb{C}$.

(1) Si g est mesurable et bornée sur Ω et si g est continue en a_0 alors

$$(\varphi_k * g)(a_0) \to g(a_0). \tag{28.286}$$

Démonstration. Pour chaque $k \in \mathbb{N}$ nous posons

$$d_k = \varphi_k * g - g. \tag{28.287}$$

Le but de ce théorème est de montrer que $d_k \to 0$ pour diverses notions de convergence.

Preuve du point (1) 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

$$d_k(a_0) = \int_{S^1}\varphi_k(s)g(a_0\bar{s})ds - \int_{S^1} g(a_0)\varphi_k(s)ds \tag{28.288a}$$

$$= \int_{S^1}\varphi_k(s)\big(g(a_0\bar{s}) - g(a_0)\big). \tag{28.288b}$$

Nous pouvons passer à la norme (et non la valeur absolue parce que d_k prend ses valeurs dans $\mathbb{C}$) :

$$|d_k(a_0)| \leqslant \int_{S^1}|\varphi_k(s)|\big|g(a_0\bar{s}) - g(a_0)\big|ds. \tag{28.289}$$

La définition d'une approximation de l'unité nous permet de considérer $M = \sup_k \|\varphi_k\|_1 < \infty$. Le lemme 28.81(2) nous permet, lui, de considérer $\alpha > 0$ tel que

$$|g(a_0\bar{s}) - g(a_0)| < \epsilon \tag{28.290}$$

dès que $s \in B(1,\alpha)$ [28]. Vu que la suite (φ_k) est une approximation de l'unité, nous avons

$$\int_{S^1\backslash B(1,\alpha)}|\varphi_k| = 0. \tag{28.291}$$

28. 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 $g(a_0\bar{s})$

Soit k suffisamment grand pour avoir $\int_{S^1\setminus B(1,\alpha)}|\varphi_k|<\epsilon$. Avec tout cela nous avons les majorations

$$|d_k(a_0)|\leqslant\int_{S^1\setminus B(1,\alpha)}|\varphi_k(s)|\big|g(a_0\bar{s})-g(a_0)\big|ds \tag{28.292a}$$

$$=\int_{S^1\setminus B(1,\alpha)}|\varphi_k(s)|\underbrace{\big|g(a_0\bar{s})-g(a_0)\big|}_{\leqslant 2\|g\|_\infty}ds+\int_{B(1,\alpha)}|\varphi_k(s)|\underbrace{\big|g(a_0\bar{s})-g(a_0)\big|}_{<\epsilon}ds \tag{28.292b}$$

$$\leqslant 2\|g\|_\infty\int_{S^1\setminus B(1,\alpha)}|\varphi_k(s)|ds+\epsilon\int_{S^1}|\varphi_k(s)|ds \tag{28.292c}$$

$$\leqslant\epsilon\big(1+2\|g\|_\infty\big). \tag{28.292d}$$

Nous avons donc bien $\lim_{k\to\infty}|d_k(a_0)|=0$ et donc la continuité de φ_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 28.96.
Soient (φ_k) une approximation de l'unité sur $\Omega=S^1$ ainsi qu'une fonction continue $g\colon S^1\to\mathbb{C}$.

(1) Si $g\in L^p(\Omega)$ $(0\leqslant p<\infty)$ et si g est continue, alors[29]

$$\varphi_k*g\xrightarrow{L^p}g. \tag{28.293}$$

(2) Si g est continue sur S^1, alors

$$\varphi_k*g\xrightarrow{L^\infty}g \tag{28.294}$$

Démonstration. En plusieurs points

$\lim_{u\to 1}\|\tau_u(g)-g\|=0$ Ceci est un petit point intermédiaire. Pour des besoins de notations, nous posons

$$\begin{aligned}\tau_u(g)\colon S^1&\to\mathbb{C}\\ s&\mapsto g(s\bar{u})\end{aligned} \tag{28.295}$$

pour $u\in S^1$.

La fonction g est continue sur le compact S^1, et y est donc uniformément continue[30]. Nous allons en déduire que $\lim_{u\to 1}\|\tau_u(g)-g\|_\infty=0$.

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 $\big|g(s)-g(a)\big|<\epsilon$. 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 28.77(2).

Pour tout $a\in S^1$ nous avons la chaîne

$$u\in B(1,\delta)\Rightarrow a\bar{u}\in B(a,\delta)\Rightarrow\big|g(a\bar{u})-g(a)\big|<\epsilon. \tag{28.296}$$

Cela étant valable pour tout a, c'est encore valable en passant au supremum[31] :

$$u\in B(1,\delta)\Rightarrow\sup_{a\in S^1}\big|g(a\bar{u})-g(a)\big|\leqslant\epsilon \tag{28.297}$$

et donc d'accord pour

$$\lim_{u\to 1}\|\tau_u(g)-g\|=0. \tag{28.298}$$

29. Vous noterez les $p\in\,]0,1[$ en bonus par rapport au cas de $\mathbb{R}^n$.
30. Théorème de Heine 13.89. C'est fondamentalement ce fait qui unifie les parties (1) et (2) de cette preuve.
31. Notez l'inégalité qui n'est plus stricte.

$\|d_k\|_\infty \to 0$ 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 τ_u, nous pouvons écrire

$$d_k(s) = \int_{S^1} \varphi_k(u) g(s\bar{u}) du - g(s) = \int_{S^1} \varphi_k(u) \big(\underbrace{g(s\bar{u})}_{=\tau_u(g)(s)} - g(s) \big) du, \tag{28.299}$$

et donc

$$|d_k(s)| \leqslant \int_{S^1} |\varphi_k(u)| \|\tau_u(g) - g\|_\infty du. \tag{28.300}$$

Nous posons $M = \sup_k \|\varphi_k\|_1$, et nous considérons δ 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 $B(1, \delta)^c$:

$$\|d_k\|_\infty \leqslant \int_{B(1,\delta)} |\varphi_k(u)| \underbrace{\|\tau_u(g) - g\|_\infty}_{\leqslant \epsilon} du + \int_{B(1,\delta)^c} |\varphi_k(u)| \|\tau_u(g) - g\|_\infty du \tag{28.301a}$$

$$\leqslant \epsilon M + 2\|g\|_\infty \int_{B(1,\delta)^c} |\varphi_k(u)| du \tag{28.301b}$$

$$\leqslant \epsilon \big(M + 2\|g\|_\infty \big) \tag{28.301c}$$

parce que pour chaque $s \in S^1$ nous avons $\tau_u(g)(s) - g(s)$ et donc $\|\tau_u(g) - g\|_\infty \leqslant 2\|g\|_\infty$.

Tout cela montre que $d_k \xrightarrow{\|.\|_\infty} 0$.

Convergence L^p, $0 < p < \infty$ Soit $\epsilon > 0$. Nous avons

$$\|d_k\|_p^p \leqslant \int_{S^1} \big| (\varphi_k * g)(s) - g(s) \big|^p ds. \tag{28.302}$$

Il existe un k à partir duquel $\|\varphi_k * g - g\|_\infty < \epsilon$. Pour de tels k nous avons

$$\|d_k\|_p^p < \epsilon^p. \tag{28.303}$$

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 13.332) nous permet de conclure que $\|d_k\|_p < \epsilon$.

Nous avons donc la convergence L^p pour $0 < p < \infty$.

Convergence L^∞ Non, la convergence L^∞ n'est pas la convergence pour la norme $\|.\|_\infty$. Voir la sous-section 28.2.2. 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

$$N_\infty(f_k - f) \leqslant \|f_k - f\| < \epsilon. \tag{28.304}$$

□

28.6.6 Base hilbertienne (suite des polynômes trigononétriques)

Voici le plan pour la suite :

— Construire un polynôme trigonométrique qui vérifie les hypothèse du lemme 28.94.

— 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 φ_k en est un.

— Invoquer le théorème 28.95(1) pour déduire que $\varphi_k * f$ est une suite de polynômes trigonométriques dans $L^2(S^1)$ qui converge $\varphi_k * f \xrightarrow{L^2} f$.

Lemme 28.97.
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 19.25 qui donne e^{ix} en termes des fonctions trigonométriques usuelles :

$$(e_1 + e_{-1})(e^{ix}) = e^{ix} + e^{-ix} = \cos(x) + i\sin(x) + \cos(x) - i\sin(x) = 2\cos(x). \tag{28.305}$$

Nous avons donc la continuité et les valeurs réelles. □

Lemme 28.98.
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 28.97 nous dit déjà que $P = e_1 + e_{-n}$ 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 [32] $M > 0$ tel que $Q = M + e_1 + e_{-1}$ est à valeurs dans $[0, \infty[$.

Une forme explicite de Q est que

$$Q(e^{ix}) = M + 2\cos(x). \tag{28.306}$$

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 28.99.
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 28.84. Il reste à prouver que ce soit un système total.

Soit $f \in L^2(S^1)$. Soit un polynôme Q vérifiant le lemme 28.98 ; nous posons

$$\varphi_k(z) = \left(\int_{S^1} Q^n \right)^{-1} Q(z)^n. \tag{28.307}$$

Cela est une approximation de l'unité par la proposition 28.94. Les φ_k sont des polynômes trigonométriques parce que les Q^n le sont et que que $\int_{S^1} Q^n$ est seulement un nombre.

Le lemme 28.93 nous dit alors que pour tout k, la fonction

$$\varphi_k * f \tag{28.308}$$

est un polynôme trigonométrique. □

Nous nous permettons de confirmer la remarque 28.86 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 $e_1 + e_{-1}$.

28.6.7 Convolution, bis

Lemme 28.100.
Nous considérons l'application

$$\begin{aligned} \varphi\colon \mathbb{R} &\to S^1 \\ t &\mapsto e^{it}. \end{aligned} \tag{28.309}$$

Soient $t, u \in \mathbb{R}$ tels que $\varphi(t) = \varphi(u)$. Alors pour toutes fonctions pour lesquelles les intégrales convergent,

$$\int_0^{2\pi} (f \circ \varphi)(\theta)(g \circ \varphi)(t - \theta)\frac{d\theta}{2\pi} = \int_0^{2\pi} (f \circ \varphi)(\theta)(g \circ \varphi)(u - \theta)\frac{d\theta}{2\pi}. \tag{28.310}$$

32. Par exemple, M est le maximum de $|P|$.

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

$$\int_0^{2\pi} (f\circ\varphi)(\theta)(g\circ\varphi)(t-\theta)\frac{d\theta}{2\pi} = \int_0^{2\pi} (f\circ\varphi)(\theta)(g\circ\varphi)(u-\theta)\frac{d\theta}{2\pi}. \tag{28.311}$$

□

Définition 28.101 (Convolution sur S^1)**.**
Le lemme 28.100 permet de définir

$$(f*g)\big(\varphi(t)\big) = \int_0^{2\pi} (f\circ\varphi)(\theta)(g\circ\varphi)(t-\theta)\frac{d\theta}{2\pi} \tag{28.312}$$

pour toutes les paires de fonctions $f, g \in \mathrm{Fun}(S^1, \mathbb{C})$ pour lesquelles l'intégrale converge.

28.7 L'espace $L^2\big([a,b]\big)$

L'espace $L^2\big([a,b]\big)$ est l'espace générique sur lequel nous allons construire les espaces L^2 sur $[-T,T]$ et $[0,2\pi]$.

Si f et g sont dans $L^2\big([a,b]\big)$, il n'est pas possible de définir $f*g$ par la formule intégrale usuelle parce que $f(x_0+t)$ n'existe pas pour tout x_0 et t dans $[a,b]$. Donc soit nous utilisons un truc pas très net comme étendre les fonctions sur $[a,b]$ en fonctions périodiques sur $\mathbb{R}$, soit nous intégrons vraiment seulement sur $[a,b]$.

Nous n'allons suivre aucune de ces deux voies ou plutôt les deux en même temps. Nous allons seulement tout ramener de S^1 que nous venons de travailler.

Définition 28.102.
Sur $[a,b]$ nous considérons la mesure de Lebesgue et le produit

$$\langle f, g\rangle = \frac{1}{b-a}\int_a^b f(x)\bar{g}(x)dx. \tag{28.313}$$

Si vous voulez une raison inavouable pour justifier ce facteur $\frac{1}{b-a}$, remarquez que dx a pour unité le mètre. En mettant la facteur $b-a$ (qui a aussi le mètre comme unité), le tout a les unités de fg, comme il se doit pour le produit scalaire.

Proposition 28.103.
L'application

$$\begin{aligned}\phi\colon L^2\big([a,b]\big) &\to L^2(S^1)\\ \phi(f)(z) &= f\big((s^{-1}\circ\varphi^{-1})(z)\big)\end{aligned} \tag{28.314}$$

où s est donnée par

$$\begin{aligned}s\colon [a,b] &\to [0,2\pi]\\ x &\mapsto 2\pi\frac{x-a}{b-a}\end{aligned} \tag{28.315}$$

et φ est l'application usuelle

$$\begin{aligned}\varphi\colon [0,2\pi[&\to S^1\\ t &\mapsto e^{it}\end{aligned} \tag{28.316}$$

est une bijection isométrique.

Démonstration. La preuve du fait que ϕ est isométrique suffira pour prouver qu'elle prend bien ses valeurs dans $L^2(S^1)$.

Isométrique C'est un calcul :

$$\begin{aligned}\|\phi(f)\|^2 &= \langle\phi(f),\phi(f)\rangle && (28.317a)\\ &= \int_{S^1}|\phi(f)|^2 && (28.317b)\\ &= \frac{1}{2\pi}\int_{[0,2\pi[}|\phi(f)\big(\varphi(u)\big)|^2du && (28.317c)\\ &= \frac{1}{2\pi}\int_0^{2\pi}|f\big((s^{-1}\circ\varphi^{-1}\circ\varphi)(u)\big)|^2du && (28.317d)\\ &= \frac{1}{2\pi}\int_0^{2\pi}|f\big(s^{-1}(u)\big)|^2du. && (28.317e)\end{aligned}$$

Il est temps de faire le changement de variables [33] $y=s^{-1}(u)$, c'est-à-dire

$$y=\frac{b-a}{2\pi}u+a. \tag{28.318}$$

En ce qui concerne la différentielle,

$$dy=\frac{b-a}{2\pi}du \tag{28.319}$$

et pour les bornes, si $u=0$ alors $y=a$ et si $u=2\pi$, $y=b$. Donc

$$\begin{aligned}\|\phi(f)\|^2 &= \frac{1}{2\pi}\int_a^b|f(y)|^2\frac{2\pi}{b-a}dy && (28.320a)\\ &= \frac{1}{b-a}\int_a^b|f|^2 && (28.320b)\\ &= \|f\|^2. && (28.320c)\end{aligned}$$

Injectif Soit f telle que $\phi(f)=0$. Alors pour tout $z\in S^1$ nous avons

$$f\big((s^{-1}\circ\varphi^{-1})(z)\big)=0. \tag{28.321}$$

Vu que $s^{-1}\circ\varphi^{-1}\colon S^1\to[a,b[$ est une bijection, pour tout $u\in[a,b[$ nous avons $f(u)=0$. Donc $f=0$ dans $L^2\big([a,b]\big)$ parce que du point de vue de L^2, que l'on prenne ou non les bornes, ce n'est pas important.

Surjectif Si $g\in L^2(S^1)$, alors en posant

$$f(u)=g\big((\varphi\circ s)(u)\big) \tag{28.322}$$

nous avons $g=\phi(f)$.

□

Définition 28.104.
En ce qui concerne le produit de convolution, si f et g sont des fonctions sur $[a,b]$ nous définissons

$$f*g=\phi^{-1}\big(\phi(f)*\phi(g)\big) \tag{28.323}$$

tant que les formules ont un sens.

33. Nous le faisons de façon un peu informelle ; soyez capable de bien justifier.

28.8 Sur $[0, 2\pi[$

Le produit de convolution est un peut subtil parce que $f(t-x)$ n'est pas défini a 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 peut toutes les opérations modulo 2π (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)

$$\begin{aligned} \varphi\colon [0, 2\pi[&\to S^1 \\ t &\mapsto e^{it} \end{aligned} \tag{28.324}$$

qui est une bijection par la proposition 19.60. Pour le produit de convolution,

$$(f * g)(x) = (f \circ \varphi^{-1}) * (g \circ \varphi^{-1})\big(\varphi(x)\big) \tag{28.325}$$

pour toutes les fonctions $f, g\colon [0, 2\pi[\to \mathbb{C}$ pour lesquelles les intégrales en jeu ont un sens.

28.9 Sur $[-T, T[$

Pour rappel, les éléments de L^2 sont des classes de fonctions à valeurs dans $\mathbb{C}$.

Proposition 28.105.
Les fonctions

$$\begin{aligned} e_n\colon [-T, T] &\to \mathbb{C} \\ t &\mapsto \frac{1}{\sqrt{2T}} e^{2\pi i n t/T}. \end{aligned} \tag{28.326}$$

forment une base hilbertienne de $L^2\big([-T, T[\big)$.

Démonstration. Nous utilisons le théorème de Stone-Weierstrass 13.303. Pour cela, nous considérons $X = [-T, T]$ qui est compact et Hausdorff ainsi que A, l'ensemble des polynômes trigonométrique :

$$A = \{ \sum_{k=-n}^{n} a_k e_k \text{ tel que } a_k \in \mathbb{C}\}_{n\in\mathbb{N}}. \tag{28.327}$$

Nous vérifions que ce A satisfait aux hypothèses de Stone-Weierstrass 13.303.

A est une algèbre Il s'agit seulement de vérifier que

$$e_k e_l = \frac{1}{\sqrt{2T}} e_{k+l}. \tag{28.328}$$

A sépare les points Soient $x, y \in [-T, T]$. Si $e_k(x) = e_k(y)$

□

28.9.1 Le cas dans $[0, 2\pi]$

En pratique, nous n'allons pas souvent travailler avec des fonctions sur intervalle symétrique $[-T, T]$, mais le plus souvent nous serons sur $[0, 2\pi]$.

Nous notons ici une conséquence du théorème 28.41 dans le cas de l'espace L^2. La proposition suivante est une petite partie du corollaire 26.49, qui sera d'ailleurs démontré de façon indépendante.

Proposition 28.106.
Si nous avons une suite de réels (a_k) telle que $\sum_{k=0}^{\infty}|a_k|^2<\infty$ alors la suite

$$f_n(x)=\sum_{k=0}^{n}a_ke^{ikx}\tag{28.329}$$

converge dans $L^2\big(]0,2\pi[\big)$.

Démonstration. Quitte à séparer les parties réelles et imaginaires, nous pouvons faire abstraction du fait que nous parlons d'une série de fonctions à valeurs dans $\mathbb{C}$ au lieu de $\mathbb{R}$.

Un simple calcul est :

$$\|f_n-f_m\|^2\leqslant\int_0^{2\pi}\sum_{k=n}^{m}|a_k|^2dx\leqslant 2\pi\sum_{k=n}^{m}|a_k|^2.\tag{28.330}$$

Par hypothèse le membre de droite est $|s_m-s_n|$ où s_k dénote la suite des sommes partielles de la série des $|a_k|^2$. Cette dernière est de Cauchy (parce que convergente dans $\mathbb{R}$) et donc la limite $n\to\infty$ (en gardant $m>n$) est zéro. Donc la suite des f_n est de Cauchy dans L^2 et donc converge dans L^2. □

Adaptons tout cela pour l'espace $L^2\big([0,2\pi]\big)$. Nous posons

$$\langle f,g\rangle=\int_0^{2\pi}f(t)\overline{g(t)}dt\tag{28.331}$$

et

$$e_n(t)=\frac{1}{\sqrt{2\pi}}e^{int}.\tag{28.332}$$

L'importance du système trigonométrique défini en 28.65 est d'être une base de $L^2\big([0,2\pi]\big)$, comme précisé dans le lemme suivant.

Lemme 28.107.
Le système trigonométrique $\{e_n\}_{n\in\mathbb{Z}}$ est une base hilbertienne de $L^2\big([0,2\pi]\big)$.

Démonstration. Pour rappel, une base hilbertienne est la définition 26.26. Nous prouvons d'abord que le système est orthogonal. Nous avons

$$\langle e_n,e_m\rangle=\frac{1}{2\pi}\int_0^{2\pi}e^{i(n-m)t}dt.\tag{28.333}$$

Si $n=m$, alors cela est égal à 1. Sinon c'est une intégrale simple :

$$\langle e_n,e_m\rangle=\left[\frac{i}{n-m}e^{i(n-m)t}\right]_0^{2\pi}=0.\tag{28.334}$$

Cela est pour l'orthogonalité.

Pour que le système soit total, il faut que son espace vectoriel engendré soit dense. Cela est le théorème 28.70. □

Note : le théorème 29.6 donné aussi la densité, mais sera démontré plus tard, indépendamment. Voir aussi les thèmes 12 et 13.

Pour un élément donné $f\in L^2\big([0,2\pi]\big)$, nous définissons

$$S_nf=\sum_{k=-n}^{n}\langle f,e_k\rangle e_k\tag{28.335}$$

et nous avons le théorème suivant, qui récompense les efforts consentis à propos de la densité des polynômes trigonométriques dans L^2.

Théorème 28.108.
Soit $f \in L^2\big([0,2\pi]\big)$. Nous avons égalité[34]

$$f = \sum_{n\in\mathbb{Z}} c_n(f) e_n \tag{28.336}$$

dans L^2.

Nous avons aussi la convergence

$$S_n f \xrightarrow{L^2} f. \tag{28.337}$$

Démonstration. Le système trigonométrique $\{e_n\}_{n\in\mathbb{Z}}$ est total pour l'espace de Hilbert $L^2\big([0,2\pi]\big)$ (sans périodicité particulière). Donc le point (1) du théorème 26.44 nous donne l'égalité demandée.

La convergence (28.337) est une reformulation de l'égalité (28.336). □

28.109.
Obtenir la convergence L^2 ne demande pas d'hypothèses de périodicité : la convergence (28.337) est automatique du fait que le système trigonométrique soit total. Ce n'est cependant pas plus qu'une convergence L^2 et elle ne demande pas $f(0) = f(2\pi)$, même si pour chacun des e_k nous avons $e_k(0) = e_k(2\pi)$.

Si $f(2\pi) \neq f(0)$, alors il existe tout de même une suite (f_n) convergente vers f au sens L^2 telle que $f_n(0) = f_n(2\pi)$. Cela ne contredit en rien le fait que $e_k(0) = e_k(2\pi)$ parce que dans L^2, la valeur d'un point seul n'a pas d'importance.

Si nous voulons une vraie convergence ponctuelle voir uniforme $(S_nf)(x) \to f(x)$, alors il faut ajouter des hypothèses sur la continuité de f, sa périodicité ou le comportement des coefficients c_n. Voir aussi le thème 6.

Exemple 28.110
Si $f \in L^2\big([0,2\pi]\big)$ est (la classe de) une fonction à valeurs réelles, alors on peut la développer avec nettement moins de termes. D'abord nous savons que $e_{-n} = \overline{e_n}$, et donc

$$\langle f, e_n\rangle = \overline{\langle f, e_{-n}\rangle}, \tag{28.338}$$

ce qui donne

$$f = \sum_{n\in\mathbb{Z}} \langle f, e_n\rangle e_n = \sum_{n\in\mathbb{N}} \langle f, e_n\rangle e_n + \overline{\langle f, e_n\rangle e_n} = \sum_{n\in\mathbb{N}} \Re\big(\langle f, e_n\rangle e_n\big). \tag{28.339}$$

Or

$$\Re\big(\langle f, e_n\rangle e_n\big) = \frac{1}{(2\pi)^{3/2}} \cos(nx) \int_0^{2\pi} f(t)\cos(nt)dt - \frac{1}{(2\pi)^{3/2}} \sin(nx) \int_0^{2\pi} f(t)\sin(nt)dt. \tag{28.340}$$

Considérons la fonction impaire $\tilde{f} \in Ł^2\big([-2\pi, 2\pi]\big)$ créée à partir de f. Elle se développe de même et nous avons la même formule (28.340) à part quelques coefficients et le fait que les intégrales sont entre -2π et 2π. Vu que $\tilde{f}$ est impaire, l'intégrale avec $\cos(nt)$ s'annule et

$$\tilde{f}(x) = \sum_{n\in\mathbb{N}} c_n \sin(nx) \tag{28.341}$$

pour certains coefficients réels c_n. Cette égalité est à considérer dans L^2, c'est-à-dire presque partout et en particulier presque partout sur $[0,2\pi]$.

Donc les fonctions réelles sur $[0,2\pi]$ peuvent être écrites sous la forme d'une série de seulement des sinus.

Note : en choisissant $\tilde{f}$ paire, nous aurions eu une série de cosinus. △

34. Notons que la somme sur $\mathbb{Z}$ dans (28.336) est commutative ; il n'est donc pas besoin d'être plus précis.

28.10 Théorème de la projection normale

28.10.1 Espace uniformément convexe

Définition 28.111 (Espace uniformément convexe[419]).
*Un espace de Banach B est **uniformément convexe** si il existe une fonction $\delta\colon]0,\infty[\to \mathbb{R}^+$ telle que si*

(1) $\|x\| \leqslant \|y\| \leqslant 1$,

(2) $\|x-y\| \geqslant \epsilon$,

alors

$$\|\frac{x+y}{2}\| \leqslant \|y\| - \delta(\epsilon). \tag{28.342}$$

Lemme 28.112 ([1]).
Si B est un espace de Banach uniformément convexe, alors pour tout $k>0$, il existe une fonction $\delta_k\colon]0,\infty[\to \mathbb{R}^2$ telle que si

(1) $\|x\| \leqslant \|y\| \leqslant k$,

(2) $\|x-y\| \geqslant \epsilon$,

alors

$$\|\frac{x+y}{2}\| \leqslant \|y\| - \delta_k(\epsilon). \tag{28.343}$$

Démonstration. Nous posons $x' = x/\|x\|$ et $y' = y/\|y\|$. Nous avons alors

$$\|x'-y'\| = \frac{\|x-y\|}{k} > \frac{\epsilon}{k}. \tag{28.344}$$

L'uniforme convexité de B dit alors que

$$\|\frac{x'+y'}{2}\| > \|y'\| - \delta(\epsilon/k). \tag{28.345}$$

En multipliant cette inégalité par k nous trouvons

$$\|\frac{x+y}{2}\| > \|y\| - k\delta(\epsilon/k). \tag{28.346}$$

Donc en posant $\delta_k(\epsilon) = k\delta(\epsilon/k)$, nous avons le résultat escompté. □

Définition 28.113 (Projection normale[419]).
Soient un espace de Banach B ainsi que $V \subset B$. Soit $a \in B$. La fonction

$$\begin{aligned} f\colon V &\to \mathbb{R} \\ x &\mapsto d(x,a) \end{aligned} \tag{28.347}$$

possède un infimum[35] *m. Si $x \in V$ est tel que $d(x,a)=m$, alors x est une **projection normale** de a sur V.*

Proposition 28.114 ([419]).
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 $\|x+y\| \leqslant \|x\| + \|y\|$ sous la forme

$$\|a - \frac{b+b'}{2}\| = \|\frac{a-b}{2} + \frac{a-b'}{2}\| \leqslant \|\frac{a-b}{2}\| + \|\frac{a-b'}{2}\| = \frac{m}{2} + \frac{m}{2} = m. \tag{28.348}$$

35. Toute fonction à valeurs positives possède un infimum, c'est la proposition 1.122.

Mais $\frac{b+b'}{2} \in V$, donc

$$\|a - \frac{b+b'}{2}\| \geqslant m. \tag{28.349}$$

Nous en déduisons que dans (28.348), toutes les inégalités sont des égalités et en particulier

$$\|\frac{b+b'}{2} - a\| = m. \tag{28.350}$$

Nous avons donc les deux égalités suivantes :

$$2m = \|a-b\| + \|a-b'\| \tag{28.351}$$

et

$$2m = \|b + b' - 2a\|. \tag{28.352}$$

Cela donne

$$\|a-b\| + \|a-b'\| = \|(a-b) + (a-b')\|. \tag{28.353}$$

Vu que B est strictement convexe, cela n'est possible que si $a - b = a - b'$, ce qui signifie que $b = b'$. □

Théorème 28.115 ([420, 419])**.**
Si B est un espace de Banach uniformément convexe, si $V \subset B$ est un sous-espace vectoriel complet et si $a \in B$, alors la projection normale de a sur V existe et est unique.

Démonstration. En deux parties.

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

$$\|b - b'\| > \epsilon > 0. \tag{28.354}$$

En posant

$$x = \frac{1}{2}\frac{b-a}{m}, \quad y = \frac{1}{2}\frac{b'-a}{m}, \tag{28.355}$$

nous avons $\|x\| = \|y\| = \frac{1}{2} < 1$. L'uniforme convexité de B donne alors

$$\|\frac{x+y}{2}\| \leqslant \|y\| - \delta(\epsilon). \tag{28.356}$$

Mais

$$x + y = \frac{\frac{1}{2}(b+b') - a}{m} \tag{28.357}$$

et

$$\|x+y\| \leqslant 2\|y\| - 2\delta(\epsilon) = 1 - 2\delta(\epsilon) < 1. \tag{28.358}$$

Nous avons donc prouvé que

$$\|\frac{1}{2}(b+b') - a\| < m, \tag{28.359}$$

ce qui est impossible parce que cela dirait que $\frac{b+b'}{2}$ est une « meilleure » projection normale que b et b'.

Existence Soient b_k dans V tels que $\|a - b_k\| \to m$. Nous supposons (quitte à passer à une sous-suite) que

$$\|a - b_{k+1}\| \leqslant \|a - b_k\|. \tag{28.360}$$

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[36] et donc complet (ce n'est pas la complétude de V qui joue ici, mais bien celle de B). Il existe $\epsilon > 0$ tel que pour tout $N \in \mathbb{N}$ il existe $p, q > N$ tels que

$$\|b_p - b_q\| > \epsilon. \tag{28.361}$$

Nous effectuons quelque choix.

36. Définition 9.24.

(1) nous choisissons $q > p$ de telle sorte que $\|a - b_p\| \leqslant \|a - b_q\|$,

(2) nous choisissons N assez grand pour avoir

$$\|a - b_p\| \leqslant \|a - b_q\| < m. \tag{28.362}$$

Nous avons $\|(a - b_p) - (a - b_q)\| = \|a_q - a_p\| > \epsilon$, ce qui avec l'uniforme convexité donne

$$\frac{\|(a - b_p) + (a - b_q)\|}{2} \leqslant \|a - b_q\| - \delta(\epsilon). \tag{28.363}$$

Donc

$$m \leqslant \|a - \frac{b_p - b_q}{2}\| = \|\frac{(a - b_p) + (a - b_q)}{2}\| \leqslant \|a - b_q\| - \delta(\epsilon) < m - \delta(\epsilon) < m. \tag{28.364}$$

Cela signifie que $m < m$, ce qui est impossible.

Conclusion La suite (b_k) converge dans B. Vu que V est complet, la limite est dans V. Cette limite, que nous nommons b, vérifie

$$\|a - b\| \leqslant \|a - b_k\| \tag{28.365}$$

pour tout k. Mais comme nous avons $m \leqslant \|a - b_k\| \to m$, nous avons $\|a - b\| = m$, c'est à dire que b est une projection normale de a sur V.

□

28.10.2 Des inégalités

Avant d'entrer dans le vif du sujet, nous nous fendons d'une petite étude de fonction. Soit

$$\begin{aligned} \phi\colon [0,1] &\to \mathbb{R} \\ x &\mapsto \frac{(1+x)^r}{1+x^r}. \end{aligned} \tag{28.366}$$

Un peu de calcul montre que

$$\frac{\phi'(x)}{\phi(x)} = \frac{r(1 - x^{r-1})}{(1 + x^r)(1 + x)}. \tag{28.367}$$

Lemme 28.116.
Soient $a, b > 0$ et $r > 1$. Nous avons les inégalités

$$a^r + b^r \leqslant (a + b)^r \leqslant 2^{r-1}(a^r + b^r). \tag{28.368}$$

Démonstration. 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

$$f'(x) = rx^{r-1} \tag{28.369a}$$

$$g'(x) = r(a + x)^{r-1}. \tag{28.369b}$$

Vu que $r > 1$, la fonction $t \mapsto t^{r-1}$ est croissante par la proposition 13.332.

Nous passons à la seconde inégalité. Le lemme 18.85 nous dit que la fonction $f\colon x \mapsto x^r$ est convexe. Donc

$$f\left(\frac{a}{2} + \frac{b}{2}\right) \leqslant \frac{1}{2}f(a) + \frac{1}{2}f(b). \tag{28.370}$$

De là nous déduisons

$$\frac{(a + b)^r}{2^r} \leqslant \frac{1}{2}(a^r + b^r), \tag{28.371}$$

c'est-à-dire la seconde inégalité. □

Nous allons démontrer les inégalités de Hanner dans le théorème 28.121. Vu que ce sera un peu longuet, nous faisons un lemme.

Lemme 28.117.
Soient $z_1, z_2 \in \mathbb{C}$. Nous avons

$$|z_1 + z_2|^p + |z_1 - z_2|^p \geqslant \big(|z_1| + |z_2|\big)^p + \big||z_1| - |z_2|\big|^p. \tag{28.372}$$

Démonstration. Soient $z_1, z_2 \in \mathbb{C}$. Nous posons

$$d = |z_1 + z_2|^p + |z_1 - z_2|^p. \tag{28.373}$$

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 [37]

$$z_2 = z_1 a^{-1} b e^{i\theta}. \tag{28.374}$$

Nous avons déjà que $z_1 + z_2 = z_1(1 + a^{-1}be^{i\theta})$ et donc

$$|z_1 + z_2| = a|1 + a^{-1}be^{i\theta}| = |a + be^{i\theta}| \tag{28.375}$$

parce que $a > 0$. De plus,

$$|a + be^{i\theta}|^2 = (a + be^{i\theta})(a + be^{-i\theta}) = a^2 + b^2 + 2ab\cos(\theta) \tag{28.376}$$

parce que $e^{i\theta} + e^{-i\theta} = \cos(\theta)$. Nous posons

$$d(\theta) = |a + be^{i\theta}|^p + |a - be^{i\theta}|^p. \tag{28.377}$$

En développant,

$$d(\theta) = \big(a^2 + b^2 + 2ab\cos(\theta)\big)^{p/2} + \big(a^2 + b^2 - 2ab\cos(\theta)\big)^{p/2}. \tag{28.378}$$

Trouvons le minimum de cette fonction de θ. D'abord sa dérivée :

$$\begin{aligned} ad'(\theta) &= pab\sin(\theta)\big[\big(a^2 + b^2 - 2ab\cos(\theta)\big)^{p/2-1} - \big(a^2 + b^2 + 2ab\cos(\theta)\big)^{p/2-1}\big] && \text{(28.379a)}\\ &= pab\sin(\theta)s(\theta). && \text{(28.379b)} \end{aligned}$$

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

$$s(0) = \big(|a - b|^2\big)^{p/2-1} - \big(|a + b|^2\big)^{p/2-1} > 0. \tag{28.380}$$

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 π. Calcul fait, nous avons

$$d(0) = d(\pi) = |a + b|^p + |a - b|^p. \tag{28.381}$$

En reliant à (28.373) nous avons l'inégalité

$$|z_1 + z_2|^p + |z_1 - z_2|^p \geqslant (a + b)^p - |a - b|^p. \tag{28.382}$$

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 (28.382) est que

$$|z_1 + z_2|^p + |z_1 - z_2|^p \geqslant \big(|z_1| + |z_2|\big)^p + \big||z_1| - |z_2|\big|^p. \tag{28.383}$$

□

37. Proposition 19.61

Encore dans la catégorie des lemmes pour les inégalités de Hanner, nous avons celui-ci.

Lemme 28.118 ([1, 421])**.**
La fonction

$$\begin{aligned}\eta\colon]0,\infty[&\to \mathbb{R}\\ a &\mapsto (a^{1/p}+1)^p+|a^{1/p}-1|^p\end{aligned} \tag{28.384}$$

est strictement convexe.

Démonstration. La fonction η est une fonction de classe C^∞ sur $]0,\infty[\backslash\{1\}$. Quelle est sa régularité en $a=1$? Le fait qu'elle y soit dérivable pas clair à cause de la valeur absolue. 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 quelque fonctions :

$$so(x)=1+x^{1/p} \tag{28.385a}$$
$$di(x)=1-x^{1/p} \tag{28.385b}$$
$$dj(x)=x^{1/p}-1 \tag{28.385c}$$

Pour les dérivées, nous avons

$$so'(x)=\frac{1}{p}x^{1/p-1} \tag{28.386a}$$
$$di'(x)=-so'(x) \tag{28.386b}$$
$$dj'(x)=so'(x). \tag{28.386c}$$

En divise les cas selon $a<1$ ou $a>1$.

Pour $a<1$ Nous avons

$$\eta(a)=so(a)^p+di(a)^p, \tag{28.387}$$

et la première dérivée donne :

$$\eta'(a)=p\,so'(a)\big(so(a)^{p-1}-di(a)^{p-1}\big). \tag{28.388}$$

Pour la seconde dérivée nous trouvons d'abord

$$\begin{aligned}\eta''(a)=&\left(\frac{1-p}{p}\right)a^{\frac{1}{p}-2}\big(so(a)^{p-1}-di(a)^{p-1}\big)\\ &+\frac{p-1}{p}a^{\frac{2}{p}-2}\big(so(a)^{p-2}+di(a)^{p-2}\big).\end{aligned} \tag{28.389}$$

À partir de là, le truc est de substituer les expressions suivantes :

$$so(a)^{p-1}=so(a)^{p-2}so(a)=so(a)^{p-2}+so(a)^{p-2}a^{1/p} \tag{28.390a}$$
$$di(a)^{p-1}=di(a)^{p-2}-x^{1/p}di(a)^{p-2}. \tag{28.390b}$$

Plein de trucs se simplifient et nous obtenons

$$\eta''(a)=\frac{p-1}{p}a^{\frac{1}{p}-2}\big(di(a)^{p-1}-so(a)^{p-2}\big). \tag{28.391}$$

Pour $a>1$ Les calculs sont essentiellement les mêmes, en partant de

$$\eta(a)=so(a)^p+dj(a)^p. \tag{28.392}$$

Les résultats sont :

$$\eta'(a)=p\,so'(a)\big(so(a)^{p-1}+dj(a)^{p-1}\big), \tag{28.393}$$

et

$$\eta''(a)=\frac{p-1}{p}a^{\frac{1}{p}-2}\big(dj(a)^{p-2}-so(a)^{p-2}\big). \tag{28.394}$$

Au final, nous avons pour tout $a \neq 1$:

$$\eta''(a) = \frac{p-1}{p} a^{\frac{1}{p}-2}\big(|1-a^{1/p}|^{p-2} - (1+a^{1/p})^{p-2}\big). \tag{28.395}$$

Ce qu'il se passe en $a = 1$ est encore une question ouvert que nous traitons maintenant.

Pour $a = 1$ Les limites des expressions (28.388) et (28.393) 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 η 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

$$f(x) = \begin{cases} di(x)^p & \text{si } x < 1 \\ dj(x)^p & \text{si } x > 1 \\ 0 & \text{si } x = 1. \end{cases} \tag{28.396}$$

Si $f'(1)$ existe, alors elle est égale à la limite

$$f'(1) = \lim_{\epsilon \to 0} \frac{f(1) - f(1-\epsilon)}{\epsilon}. \tag{28.397}$$

Les deux limites à calculer sont :

$$\lim_{\epsilon \to 0^+} \frac{\big((1+\epsilon)^{1/p} - 1\big)^p}{\epsilon} \tag{28.398}$$

et

$$\lim_{\epsilon \to 0^-} \frac{\big(1 - (1+\epsilon)^{1/p}\big)^p}{\epsilon}. \tag{28.399}$$

La première se traite par la règle de l'Hospital [38], et le résultat est zéro. Pour la seconde, il faut juste transformer

$$\lim_{\epsilon \to 0^+} \frac{\big((1+\epsilon)^{1/p} - 1\big)^p}{\epsilon} = \lim_{h \to 0^+} \frac{\big(1 - (1-h)^{1/p}\big)^p}{-h}, \tag{28.400}$$

qui se traite également par la règle de l'Hospital. Le résultat est également zéro.

Donc η est dérivable en $a = 1$ et la dérivée vaut $\eta'(1) = 2^{p-1}$.

En récapitulant, nous avons $\eta'' > 0$ sur $]0, \infty[\backslash\{0\}$, donc η' est croissante sur cette partie (proposition 13.124). Vu que η' est continue sur $]0, \infty[$, elle est même croissante (strictement) sur tout $]0, \infty[$.

La proposition 18.81 conclu que η est strictement convexe sur $]0, \infty[$. □

Toujours dans la catégorie des lemmes pour les inégalités de Hanner, nous avons celui-ci.

Lemme 28.119 ([421])**.**
Soit $1 < p < 2$. La fonction

$$\begin{aligned} \xi\colon \mathbb{R}^+ \times \mathbb{R}^+ &\to \mathbb{R} \\ (a,b) &\mapsto \big(a^{1/p} + b^{1/p}\big)^p + |a^{1/p} - b^{1/p}|^p \end{aligned} \tag{28.401}$$

est convexe.

Pour rappel, les conventions de données en 1.101 donnent $\mathbb{R}^+ = [0, \infty[$.

Démonstration. La fonction ξ vérifie facilement les conditions suivante :

— $\xi(a,b) = \xi(b,a)$,

— $\xi(0,0) = 0$,

38. Proposition 13.131

— $\xi(ta,tb)=t\xi(a,b)$ pour tout $t\geqslant 0$.

Nous posons

$$\begin{aligned}\eta\colon \mathbb{R}^+ &\to \mathbb{R}\\ a &\mapsto \xi(a,1).\end{aligned}\tag{28.402}$$

Le lemme 28.118 dit que η est strictement convexe, et le lemme 18.91 conclu que ξ est convexe. □

Lemme 28.120 ([422]).
Soit $1<p<2$. Nous considérons les fonctions

$$\begin{aligned}\alpha\colon [0,1] &\to \mathbb{R}\\ x &\mapsto (1+x)^{p-1}+(1-x)^{p-1}\end{aligned}\tag{28.403}$$

et

$$\begin{aligned}\beta\colon [0,1] &\to \mathbb{R}\\ x &\mapsto x^{1-p}\big((1+x)^{p-1}-(1-x)^{p-1}\big).\end{aligned}\tag{28.404}$$

Soient $A,B\in\mathbb{R}$. Nous avons

$$\alpha(x)|A|^p+\beta(x)|B|^p\leqslant |A+B|^p+|A-B|^p.\tag{28.405}$$

Démonstration. Plusieurs étapes.

$\beta(x)\leqslant\alpha(x)$ 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

$$\alpha'(x)=(p-1)(x+1)^{p-2}-(p-1)(1-x)^{p-2}.\tag{28.406}$$

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,

$$\beta'(x)=\frac{p-1}{x^p}\big((1-x)^{p-2}-(1+x)^{p-2}\big).\tag{28.407}$$

Cela nous permet de calculer $\alpha'-\beta'$:

$$\alpha'(x)-\beta'(x)=(p-1)\big(1+\frac{1}{x^p}\big)\big((1+x)^{p-2}-(1-x)^{p-2}\big).\tag{28.408}$$

Vu que $1<p<2$, le nombre $p-2$ est strictement négatif ; afin de travailler avec des exposants positifs, nous écrivons

$$\alpha'(x)-\beta'(x)=\underbrace{(p-1)}_{>0}\underbrace{\big(1+\frac{1}{x^p}\big)}_{>0}\underbrace{\left(\frac{1}{(1+x)^{2-p}}-\frac{1}{(1-x)^{2-p}}\right)}_{<0}.\tag{28.409}$$

Nous avons $\alpha'(x)-\beta'(x)<0$ pour tout $x\in\,]0,1]$. Du fait qu'en plus nous ayons $\alpha(1)=\beta(1)$, nous déduisons que $\alpha(x)\geqslant\beta(x)$.

Une petite étude de fonction Soit $R\in[0,1]$. Nous considérons la fonction

$$\begin{aligned}F\colon [0,1]\mathbb{R} &\to \mathbb{R}\\ x &\mapsto \alpha(x)+R^p\beta(x).\end{aligned}\tag{28.410}$$

Nous montrons maintenant que cette fonction a un maximum global pour $x=R$. D'abord sa dérivée :

$$F'(x)=\underbrace{(p-1)}_{>0}\underbrace{\big((1-x)^{p-1}-(1+x)^{p-2}\big)}_{<0}\left(1-\left(\frac{R}{x}\right)^p\right)\tag{28.411}$$

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.

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.

Pour $0 < B < A$ Nous devons prouver que

$$\alpha(x)A^p + \beta(x)B^p \leqslant (A+B)^p + (A-B)^p. \tag{28.412}$$

En divisant par A^p et en posant $R = B/A$, l'inéquation (28.412) est équivalente à

$$\alpha(x) + \beta(x)R^p \leqslant (1+R)^p + (1-R)^p \tag{28.413}$$

où $R \in]0,1[$ parce que nous avons supposé $0 < B < A$. Nous avons (il y a un petit calcul pour $F(R)$)

$$(1+R)^p + (1-R)^p = F(R) \geqslant F(x) = \alpha(x) + \beta(x)R^p. \tag{28.414}$$

ok.

Pour $0 < A < B$ Lorsque $0 < A < B$ nous avons

$$\begin{aligned}
\alpha(x)|A|^p + \beta(x)|B|^p &= \alpha(x)A^p + \beta(x)^p && \text{(28.415a)}\\
&\leqslant \alpha(x)B^p + \beta(x)A^p && \text{(28.415b)}\\
&\leqslant (B+A)^p + (B-A)^p && \text{(28.415c)}\\
&= |A+B|^p + |A-B|^p. && \text{(28.415d)}
\end{aligned}$$

Justification :

— Pour (28.415b), c'est parce que $\alpha(x) > \beta(x)$; alors en mettant le plus grand de A et B devant le α au lieu du β, nous majorons.

— Pour (28.415c), c'est l'inégalité dans le cas $0 < B < A$, mais en inversant les noms de A et B.

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 $A = B$.

Pour $A < 0$, $B > 0$ En posant $A' = -A$ nous avons $A' > 0$ et nous pouvons écrire

$$|A+B|^p + |A-B|^p = |-A'+B|^p + |-A'-B|^p = |B-A'|^p + |B+A'|^p \geqslant \alpha(x)|A'|^p + \beta(x)|B|^p. \tag{28.416}$$

Nous avons utilisé, avec A' et B le cas déjà prouvé $A', B > 0$.

Pour $A > 0$, $B < 0$ Celui-là, je vous le laisse.

Pour $A < 0$, $B < 0$ Posez $A' = -A$ et $B' = -B$ et hop.

□

Théorème 28.121 (Inégalités de Hanner[421, 422])**.**
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$. Soit $1 < p < 2$ and $f, g \in L^p(\Omega, \mathcal{A}, \mu)$; nous avons

$$\big(\|f\|_p + \|g\|_p\big)^p + \Big|\|f\|_p - \|g\|_p\Big|^p \leqslant \|f+g\|_p^p + \|f-g\|_p^p \tag{28.417}$$

Il y a égalité si et seulement si $f(t)$ et $g(t)$ sont colinéaires pour presque tout t.

Démonstration. Nous supposons que $\|f\|_p \geqslant \|g\|_p$ 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 28.120 pour $A = |f(\omega)|$ et $B = |g(\omega)|$:

$$\alpha(x)|f(\omega)|^p + \beta(x)|g(\omega)|^p \leqslant \big|f(\omega) + g(\omega)\big|^p + \big|f(\omega) - g(\omega)\big|^p. \tag{28.418}$$

Nous intégrons cela par rapport à ω sur Ω :

$$\alpha(x)\|f\|_p^p + \beta(x)\|g\|_p^p \leqslant \|f+g\|_p^p + \|f-g\|_p^p. \tag{28.419}$$

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 $x = \frac{\|g\|_p}{\|f\|_p}$, ce qui est permis parce que nous avons supposé $\|f\|_p \geqslant \|g\|_p$. Faites le calcul, collectez les termes identiques, vous obtiendrez

$$\big(\|f\|_p + \|g\|_p\big)^p + \big(\|f\|_p - \|g\|_p\big)^p \leqslant \|f+g\|_p^p + \|f-g\|_p^p. \tag{28.420}$$

Et vu que $\|f\|_p \geqslant \|g\|_p$, nous pouvons gratuitement faire

$$\|f\|_p - \|g\|_p = \big|\|f\|_p - \|g\|_p\big|. \tag{28.421}$$

Fini pour Hanner. □

28.10.3 Inégalités de Clarkson

Lemme 28.122 ([423])**.**
Si $p \geqslant 2$ et si $a, b \in \mathbb{C}$, alors

$$\left|\frac{a+b}{2}\right|^p + \left|\frac{a-b}{2}\right|^p \leqslant \frac{1}{2}\big(|a|^p + |b|^p\big). \tag{28.422}$$

Démonstration. Nous prouvons l'inégalité en prenant montent petit à petit en généralité.

Avec $x > 0$ Soit $x \geqslant 0$. Nous montrons dans cette partie l'inégalité

$$x^p + 1 \leqslant (x+1)^{p/2}. \tag{28.423}$$

Pour cela nous considérons la fonction

$$\begin{aligned} f\colon [0,\infty[&\to \mathbb{R} \\ t &\mapsto (t^2+1)^{p/2} - t^p - 1. \end{aligned} \tag{28.424}$$

Nous avons $f(0) = 0$, mais aussi, en utilisant les règle de dérivation [39] nous trouvons vite

$$f'(t) = p(t^2+1)^{p/2-1}t - pt^{p-1}. \tag{28.425}$$

Vu que $(t^2+1)^{p/2-1} \geqslant t^{p-2}$, le signe de $f'(t)$ est toujours strictement positif pour $t > 0$. La proposition 13.124 fait que f est strictement croissante et que $f(t) > 0$ pour tout $t > 0$.

Avec $x, y \geqslant 0$ Soient $x, y \geqslant 0$ dans $\mathbb{R}$. Nous prouvons dans cette partie que

$$(x^2+y^2)^{p/2} \geqslant x^p + y^p. \tag{28.426}$$

Il s'agit d'appliquer l'inégalité (28.423) à x/y :

$$\left(\left(\frac{x}{y}\right)^2 + 1\right)^{p/2} \geqslant \left(\frac{x}{y}\right)^p + 1. \tag{28.427}$$

En multipliant par y^p et en simplifiant un peu, nous trouvons le résultat (28.426).

39. Par exemple celle de la proposition 15.242.

Avec $a,b\in\mathbb{C}$ Nous appliquons l'inégalité (28.426) à $x=|\frac{a+b}{2}|$ et $y=|\frac{a-b}{2}|$. Cela donne :

$$\left|\frac{a+b}{2}\right|^p+\left|\frac{a-b}{2}\right|^p\leqslant\left(\left|\frac{a+b}{2}\right|^2+\left|\frac{a-b}{2}\right|^2\right)^{p/2} \tag{28.428a}$$

$$=\left(\frac{2|a|^2+2|b|^2}{4}\right)^{p/2} \tag{28.428b}$$

$$=\frac{1}{2}|a|^p+\frac{1}{2}|b|^p. \tag{28.428c}$$

□

Lemme 28.123 ([424]).
Soient $a,b\in\mathbb{C}$ ainsi que $1<p<2$. Nous notons q l'exposant conjugué de p. Nous avons l'inégalité

$$|a+b|^q+|a-b|^q\leqslant 2\big(|a|^p+|b|^p\big)^{q-1}. \tag{28.429}$$

Démonstration. Soient $a,b\in\mathbb{C}$; nous nommons $|a|$ le module du nombre complexe a. Nous avons :

$$|a+b|^q+|a-b|^q=\left\|\begin{pmatrix}|a+b|\\|a-b|\end{pmatrix}\right\|_q^q \tag{28.430a}$$

$$\leqslant\left(2^{\frac{1}{q}-\frac{1}{2}}\left\|\begin{pmatrix}|a+b|\\|a-b|\end{pmatrix}\right\|_2\right)^q \tag{28.430b}$$

$$=2^{1-q/2}\left\|\begin{pmatrix}|a+b|\\|a-b|\end{pmatrix}\right\|_2^q \tag{28.430c}$$

$$=2^{1-q/2}\big(|a|^2+|b|^2\big)^{q/2} \tag{28.430d}$$

$$=2\left\|\begin{pmatrix}|a|\\|b|\end{pmatrix}\right\|_2^q \tag{28.430e}$$

$$\leqslant 2\left\|\begin{pmatrix}|a|\\|b|\end{pmatrix}\right\|_p^q \tag{28.430f}$$

$$=2\big(|a|^p+|b|^p\big)^{q/p} \tag{28.430g}$$

$$=2\big(|a|^p+|b|^p\big)^{q-1}. \tag{28.430h}$$

Quelques justifications.

— Pour (28.430b). C'est Hölder de la proposition 18.106.

— Pour (28.430d). C'est le calcul suivant, basé sur le fait que $|z|^2=z\bar{z}$: Un petit calcul pour la norme 2 :

$$\left\|\begin{pmatrix}|a+b|\\|a-b|\end{pmatrix}\right\|_2^2=|a+b|+|a-b|^2 \tag{28.431a}$$

$$=|a|^2+a\bar{b}+\bar{a}b+|b|^2 \tag{28.431b}$$

$$+|a|^2-a\bar{b}-\bar{a}b+|b|^2 \tag{28.431c}$$

$$=2|a|^2+2|b|^2. \tag{28.431d}$$

— Pour (28.430f). Du fait que $p<2<q$, la proposition 18.106 donne $\|x\|_p\geqslant\|x\|_2$.

— Pour (28.430h). Simplement multiplier par q l'équation $\frac{1}{p}+\frac{1}{q}=1$ fournit $\frac{q}{p}=q-1$.

□

Proposition 28.124 (Inégalité de Clarkson[420]).
Soient $f,g\in L^p(\Omega,\mathcal{A},\mu)$.

(1) Si $p \geqslant 2$, alors

$$\|\frac{f+g}{2}\|_p^p + \|\frac{f-g}{2}\|_p^p \leqslant \frac{1}{2}\Big(\|f\|_p^p + \|g\|_p^p\Big). \tag{28.432}$$

(2) Si $1 < p < 2$ et si q est l'exposant conjugué de p, alors

$$\|f+g\|_p^q + \|f-g\|_p^q \leqslant 2\Big(\|f\|_p^p + \|g\|_p^p\Big)^{q-1}, \tag{28.433}$$

ou

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f-g}{2}\|_p^q \leqslant 2^{1-q}\big(\|f\|_p^p + \|g\|_p^p\big)^{q-1}. \tag{28.434}$$

Démonstration. En deux parties.

Pour $p \geqslant 2$ 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 28.122 :

$$\left|\frac{f(\omega)+g(\omega)}{2}\right|^p + \left|\frac{f(\omega)-g(\omega)}{2}\right|^p \leqslant \frac{1}{2}\big(|f(\omega)|^p + |g(\omega)|^p\big). \tag{28.435}$$

Nous avons les substitutions évidentes $f(\omega)+g(\omega) = (f+g)(\omega)$ et $f(\omega)-g(\omega) = (f-g)(\omega)$. En intégrant alors (28.435) sur Ω nous trouvons l'inégalité demandée.

Pour $1 < p < 2$ Il s'agit de faire la même chose, en utilisant l'inégalité de Clarkson du lemme 28.123.

Pour obtenir (28.434), il s'agit simplement de multiplier et diviser le member de gauche de (28.433) par 2^q.

□

28.10.4 Uniforme convexité des espaces de Lebesgue

Proposition 28.125 ([425]).
Si $1 < p < \infty$, l'espace $L^p(\Omega, \mathcal{A}, \mu)$ est uniformément convexe[40].

Démonstration. En deux parties.

$1 < p \leqslant 2$ Nous montrons que la fonction $\delta(\epsilon) = 2^{-q}\epsilon^q$ fonctionne.

Soient $f, g \in L^p$ telles que $\|f\|_p \leqslant \|g\|_p \leqslant 1$ et $\|f-g\|_p \geqslant \epsilon$. Nous commençons par écrire l'inégalité de Clarkson (28.433) :

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f-g}{2}\|_p^q \leqslant 2^{1-q}\big(\|f\|_p^p + \|g\|_p^p\big)^{q-1}. \tag{28.436}$$

Par hypothèse, $\|f\|_p$ et $\|g\|_p$ sont plus petites que 1. Vu que $p > 1$, nous avons

$$\|f\|_p^p + \|g\|_p^p \leqslant 1 + 1 = 2. \tag{28.437}$$

En remplaçant dans le membre de droite de (28.436) nous avons

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f-g}{2}\|_p^q \leqslant 2^{1-q}2^{q-1} = 1, \tag{28.438}$$

et donc

$$\|\frac{f+g}{2}\|_p^q \leqslant 1 - \|\frac{f-g}{2}\|_p^q. \tag{28.439}$$

Par ailleurs nous avons supposé $\|f-g\|_p \geqslant 1$. Donc aussi[41]

$$\|\frac{f-g}{2}\|_p^q \geqslant 2^{-q}\epsilon^q. \tag{28.440}$$

40. Définition 28.111.
41. Ici j'ai un coefficient un peu différent que celui de [425]. Écrivez-moi pour confirmer ou infirmer mes calculs.

Et par un autre ailleurs,

$$\|\frac{f+g}{2}\|_p = \frac{1}{2}\|f+g\|_p \leqslant \frac{1}{2}\big(\|f\|_p + \|g\|_p\big) \leqslant 1. \tag{28.441}$$

Vu que nous avons $q \geqslant 2$, cela donne aussi

$$\|\frac{f+g}{2}\|_p \leqslant \|\frac{f+g}{2}\|^q. \tag{28.442}$$

Avec les inégalités (28.440) et 28.442 nous finissons l'inégalité (28.439) :

$$\|\frac{f+g}{2}\|_p \leqslant \|\frac{f+g}{2}\|_p^q \leqslant 1 - 2^{-q}\epsilon^q \leqslant \|g\|_p - \delta(\epsilon). \tag{28.443}$$

Okay, c'est bon.

$2 \leqslant p < \infty$ Il s'agit de faire la même chose en partant de Clarkson (28.432). Le résultat est que la fonction $\delta(\epsilon) = (\sigma/2)^p$, ça fonctionne.

□

28.10.5 Théorème de la projection normale

Proposition 28.126.
Si $1 < p < \infty$, et si V est un sous-espace vectoriel fermé de $L^p(\Omega, \mathcal{A}, \mu)$, alors la projection normale[42] *de $a \in L^p$ sur V existe et est unique.*

Démonstration. La proposition 28.125 nous indique que l'espace $L^p(\Omega, \mathcal{A}, \mu)$ est uniformément convexe. Or le théorème 28.115 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 $p \geqslant 2$.

Théorème 28.127 (Théorème de la projection normale[426])**.**
Nous considérons $p \geqslant 2$. Soit un sous-espace vectoriel fermé $W \subset L^p(\Omega, \mathcal{A}, \mu)$ et $u_0 \in L^p$. Nous notons

$$d(u_0, W) = \inf_{w \in W} d(u_0, W). \tag{28.444}$$

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 $p > 2$.

$p = 2$ Pour $p = 2$, nous savons que L^2 est un espace de Hilbert[43], et nous avons déjà le théorème de la projection 26.5.

$p > 2$ Pour chaque $x \in \Omega$ nous avons $f(x), g(x) \in \mathbb{C}$ et donc l'identité du parallélogramme[44] :

$$\big|f(x) - g(x)\big|^2 + \big|f(x) + g(x)\big| = 2|f(x)| + 2|g(x)|^2. \tag{28.445}$$

Vu que $p > 2$, la fonction $s\colon x \mapsto x^{p/2}$ est convexe (lemme 18.85). Calcul :

$$\begin{aligned}
|f(x) - g(x)|^p + |f(x) + g(x)|^p &= \big(|f(x) - g(x)|^2\big)^{p/2} + \big(|f(x) + g(x)|^2\big)^{p/2} && \text{(28.446a)}\\
&= s\big(|\dots|^2\big) + s\big(|\dots|^2\big) && \text{(28.446b)}\\
&\leqslant \big(|f(x) - g(x)|^2 + |f(x) + g(x)|^2\big)^{p/2} && \text{(28.446c)}\\
&= \big(2|f(x)|^2 + 2|g(x)|^2\big)^{p/2} && \text{(28.446d)}\\
&= 2^{p/2}\big(|f(x)|^2 + |g(x)|^2\big)^{p/2} && \text{(28.446e)}\\
&\leqslant 2^{p/2}2^{p/2-1}\big(|f(x)|^p + |g(x)|^p\big) && \text{(28.446f)}\\
&= 2^{p-1}\big(|f(x)|^p + |g(x)|^p\big) && \text{(28.446g)}
\end{aligned}$$

42. Définition 28.113.
43. Lemme 28.73.
44. Proposition 11.11 en remarquant que $(z_1, z_2) \mapsto z_1\bar{z}_2$ est un produit scalaire hermitien sur $\mathbb{C}$.

Justifications :

— Pour (28.446c) : la convexité de s.

— Pour (28.446d) : la relation (28.445).

— Pour (28.446f) : la seconde inégalité du lemme 28.446f.

Nous isolons $|f(x)-g(x)|^p$:

$$|f(x)-g(x)|^p \leqslant 2^{p-1}\big(|f(x)|^p+|g(x)|^p\big)-|f(x)+g(x)|^p \tag{28.447a}$$

$$= 2^p\left(\frac{|f(x)|^p+|g(x)|^p}{2}-\left|\frac{|f(x)|+|g(x)|}{2}\right|^p\right) \tag{28.447b}$$

Cette inégalité étant valable pour tout x, nous pouvons intégrer sur Ω et découper l'intégrale en petits morceaux :

$$\|f-g\|_p^p \leqslant 2^p\left(\frac{\|f\|_p^p+\|g\|_p^p}{2}-\|\frac{f+g}{2}\|_p^p\right). \tag{28.448}$$

Voila une bonne chose de prouvée. Nous pouvons maintenant passer au vif du sujet.

Soit une suite w_j dans W telle que $\|u_0-w_j\|\to d(u_0,W)$. Trois choses à savoir sur cette suite :

(1) Une telle suite existe parce que $d(u_0,W)$ est défini comme un infimum.

(2) Rien ne garanti qu'elle converge.

(3) 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

$$\|\frac{w_j+w_k}{2}-u_0\|\geqslant d(u_0,W). \tag{28.449}$$

En tenant compte de cela, nous écrivons l'inégalité (28.448) avec $f=w_j-u_0$ et $g=w_k-u_0$:

$$\|f-g\|_p^p=\|w_j-w_k\|_p^p\leqslant 2^p\left(\frac{\|w_j-u_0\|^p+\|w_k-u_0\|^p}{2}-d(u_0,W)\right). \tag{28.450}$$

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 $\|w_j-u_0\|^p\leqslant d(u_0,W)^p+\epsilon_1$ et $\|w_k-u_0\|^p\leqslant d(u_0,W)^p+\epsilon_2$. Pour de telles valeurs de j et k, nous avons

$$\|w_j-w_k\|_p\leqslant 2\left(\frac{\epsilon_1+\epsilon_2}{2}\right)<2\epsilon^{1/p}. \tag{28.451}$$

Donc la suite (w_j) est de Cauchy.

L'espace L^p étant complet par le théorème 28.40, 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

$$\|w-u_0\|=\lim_j\|w_j-u_0\|=d(W,u_0). \tag{28.452}$$

□

28.11 Dualité 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.

Voici déjà un bel énoncé. Pour des espaces mesurés $(\Omega, \mathcal{A}, \mu)$ plus généraux, voir l'arme totale en el théorème 28.130.

Proposition 28.128 ([168], thème 33).
Soit $1 < p < 2$ et q tel que $\frac{1}{p} + \frac{1}{q} = 1$. L'application

$$\begin{aligned} \Phi\colon L^q\big([0,1]\big) &\to L^p\big([0,1]\big)' \\ \Phi_g(f) &= \int_{[0,1]} f\bar{g}. \end{aligned} \tag{28.453}$$

est une isométrie linéaire surjective.

Démonstration. Pour la simplicité des notations nous allons noter L^2 pour $L^2\big([0,1]\big)$, et pareillement pour L^p.

Φ_g est un élément de $(L^p)'$ Si $f \in L^p$ et $g \in L^q$ nous devons prouver que $\Phi_q(f)$ est bien définie. Pour cela nous utilisons l'inégalité de Hölder[45] qui dit que $fg \in L^1$; par conséquent la fonction $f\bar{g}$ est également dans L^1 et nous avons

$$|\Phi_g(f)| \leqslant \int_{[0,1]} |f\bar{g}| = \|fg\|_1 \leqslant \|f\|_p \|g\|_q. \tag{28.454}$$

En ce qui concerne la norme de l'application Φ_g nous avons tout de suite

$$\|\Phi_g\| = \sup_{\|f\|_p=1} \big|\Phi_g(f)\big| \leqslant \|g\|_q. \tag{28.455}$$

Cela signifie que l'application Φ_g est bornée et donc continue par la proposition 12.25. Nous avons donc bien $\Phi_g \in (L^p)'$.

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 $|\Phi_g| \geqslant \|g\|_q$, ce qui conclurait que $|\Phi_g| = \|g\|_q$.

Nous posons $f = g|g|^{q-2}$, de telle sorte que $|f| = |g|^{q-1}$ et

$$\|f\|_p = \left(\int |g|^{p(q-1)} \right)^{1/p} = \left(\int |g|^q \right)^{1/p} = \|g\|_q^{q/p} \tag{28.456}$$

où nous avons utilisé le fait que $p(q-1) = q$. La fonction f est donc bien dans L^p. D'autre part,

$$\Phi_g(f) = \int f\bar{g} = \int g|g|^{q-2}\bar{g} = \int |g|^q = \|g\|_q^q. \tag{28.457}$$

Donc

$$\frac{|\Phi_g(f)|}{\|f\|_p} = \|g\|_q^{q-\frac{q}{p}} = \|g\|_q \tag{28.458}$$

où nous avons encore utilisé le fait que $q - \frac{q}{p} = \frac{q(p-1)}{p} = 1$.

Surjectif Soit $\ell \in (L^p)'$; c'est une application $\ell\colon L^p \to \mathbb{C}$ dont nous pouvons prendre la restriction à L^2 parce que la proposition 28.38 nous indique que $L^2 \subset L^p$. Nous nommons $\phi\colon L^2 \to \mathbb{C}$ cette restriction.

45. Proposition 28.33.

$\phi \in (L^2)'$ Nous devons montrer que ϕ 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 :

$$\sup_{f\in L^2} \frac{|\phi(f)|}{\|f\|_2} \leqslant \sup_{f\in L^2} \frac{|\ell(f)|}{\|f\|_p} < \infty. \tag{28.459}$$

Nous avons utilisé l'inégalité de norme $\|f\|_p \leqslant \|f\|_2$ de la proposition 28.38(2).

Utilisation du dual de L^2 Étant donné que L^2 est un espace de Hilbert (lemme 28.73) et que $\phi \in (L^2)'$, le théorème 26.17 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 $\ell(f) = \Phi_g(f)$.

$g \in L^q$ Nous posons $f_n = g|g|^{q-2}\mathbb{1}_{|g|<n}$. Nous avons d'une part

$$\Phi_g(f_n) = \int_0^1 f_n \bar{g} = \int_{|g|<n} |g|^q. \tag{28.460}$$

Et d'autre part comme $f_n \in L^2$ nous avons aussi $\phi(f_n) = \Phi_g(f_n)$ et donc

$$0 \leqslant \Phi(f_n) = \phi(f_n) \leqslant \|\ell\|\|f_n\|_p \tag{28.461a}$$

$$= \|\ell\| \left(\int_{|g|<n} |g|^{(q-1)p} \right)^{1/p} \tag{28.461b}$$

$$= \|\ell\| \left(\int_{|g|<n} |g|^q \right)^{1/p}. \tag{28.461c}$$

où nous avons à nouveau tenu compte du fait que $p(q-1) = q$. En combinant avec (28.460) nous trouvons

$$\int_{|g|<n} |g|^q \leqslant \|\ell\| \left(\int_{|g|<n} |g|^q \right)^{1/p}, \tag{28.462}$$

et donc

$$\left(\int_{|g|<n} |g|^q \right)^{1-\frac{1}{p}} \leqslant \|\ell\|, \tag{28.463}$$

c'est-à-dire

$$\Big(\int_{|g|<n} |g|^q \Big)^{1/q} \leqslant \|\ell\|. \tag{28.464}$$

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

$$g_n = |g|^q \mathbb{1}_{|g|<n} \tag{28.465}$$

qui est une suite croissante de fonctions convergeant ponctuellement vers $|g|^q$. Le théorème de Beppo-Levi 15.160 nous permet alors d'écrire

$$\lim_{n\to\infty} \int_{|q|<n} |g|^q = \int_0^1 |g|^q. \tag{28.466}$$

Mais comme pour chaque n nous avons $\int_{|g|<n} |g|^q \leqslant \|\ell\|^q$, nous conservons l'inégalité à la limite et

$$\int_0^1 |g|^q \leqslant \|\ell\|^q. \tag{28.467}$$

Cela prouve que $g \in L^p$.

$\underline{\ell(f) = \Phi_g(f)}$ Soit $f \in L^p$. En vertu de la densité de L^2 dans L^p prouvée dans le corollaire 28.48 nous pouvons considérer une suite (f_n) dans L^2 telle que $f_n \xrightarrow{L^p} f$. Pour tout n nous avons

$$\ell(f_n) = \Phi_g(f_n). \tag{28.468}$$

Mais ℓ et Φ_g étant continues sur L^p nous pouvons prendre la limite et obtenir

$$\ell(f) = \Phi_g(f). \tag{28.469}$$

□

Lemme 28.129 ([6]).
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

$$\frac{1}{\mu(E)} \int_E g d\mu \in S, \tag{28.470}$$

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 [46], nous aurions $g(x) \in S^c$ presque nulle part.

Vu que $\frac{1}{\mu(E)} \int_E a = a$ nous avons

$$\left|\frac{1}{\mu(E)} g d\mu - a\right| = \left|\frac{1}{\mu(E)} \int_E (g - a)\right| \leqslant \frac{1}{\mu(E)} \int_E |g - a| \leqslant \frac{1}{\mu(E)} \mu(E) r = r. \tag{28.471a}$$

Donc

$$\frac{1}{\mu(E)} \int_E g d\mu \in D, \tag{28.472}$$

ce qui est une contradiction avec le fait que $D \subset S^c$. □

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 $\mathbb{C}$.

Théorème 28.130 (Théorème de représentation de Riesz, thème 33, [6, 261, 427, 250]).
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

$$\begin{aligned} \Phi &: L^q \to (L^p)' \\ \Phi_g(f) &= \int_\Omega f \bar{g} d\mu \end{aligned} \tag{28.473}$$

est une bijection isométrique 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 σ-fini.

Démonstration. Par petits bouts.

$\underline{\textbf{$\Phi$ est injective}}$ Nous commençons par prouver que Φ est injectif. Soient $g, g' \in L^q$ tels que $\Phi_g = \Phi_{g'}$. Alors pour tout $f \in L^p$ nous avons

$$\int_\Omega f(g - g') d\mu = 0. \tag{28.474}$$

46. Tout ouvert peut être écrit comme union dénombrable d'éléments d'une base de topologie par la proposition 7.55 et $\mathbb{C}$ a une base dénombrable de topologie par la proposition 7.91.

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

$$\int_{A_i} (g - g')d\mu = \int_{\Omega} \mathbb{1}_{A_i}(x)(g - g')(x)d\mu(x) = 0. \tag{28.475}$$

La proposition 28.22 nous dit alors que $g - g' = 0$ dans $L^q(A_i)$. Pour chaque i, la partie $N_i = \{x \in A_i \text{ tel que } (g - g')(x) = 0\}$ est de mesure nulle.

Vu que Ω est l'union de tous les A_i, la partie de Ω 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 Ω, ce qui signifie $g - g' = 0$ dans $L^q(\Omega)$.

La suite La partie difficile est de montrer que Φ est surjective.

Soit $\phi \in L^p(\Omega)'$. Si $\phi = 0$, c'est bien dans l'image de Φ ; 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.

Une mesure complexe Si $E \in \mathcal{A}$ nous notons $\nu(E) = \phi(\mathbb{1}_E)$. Nous prouvons maintenant que ν est une mesure complexe[47] sur $(\Omega, \mathcal{A})$. La seule condition pas facile est la condition de dénombrable additive. 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 $E_k = \bigcup_{i \leqslant k} A_i$ pour avoir $\bigcup_k A_k = \bigcup_k E_k$ avec l'avantage que les E_k soient emboîtés. Cela donne

$$\|\mathbb{1}_E - \mathbb{1}_{E_k}\|_p = \mu(E \backslash E_k)^{1/p}, \tag{28.476}$$

mais vu que $1 \leqslant p < \infty$, avoir $x_k \to 0$ implique d'avoir $x_k^{1/p} \to 0$. Prouvons que $\mu(E \backslash E_k) \to 0$. En vertu du lemme 15.22 nous avons pour chaque k :

$$\mu(E \backslash E_k) = \mu(E) - \mu(E_k), \tag{28.477}$$

et vu que $E_k \to E$ est une suite croissante, le lemme 15.23(1), sachant que μ est une mesure « normale », donne

$$\lim_{n \to \infty} \mu(E_k) = \mu\Big(\bigcup_k E_k\Big). \tag{28.478}$$

Donc effectivement $\mu(E_k) \to \mu(E)$ et donc oui : $\mu(E \backslash E_k) \to 0$. Jusqu'à présent nous avons

$$\lim_{k \to \infty} \|\mathbb{1}_E - \mathbb{1}_{E_k}\|_p = 0, \tag{28.479}$$

c'est-à-dire $\mathbb{1}_{E_k} \xrightarrow{L^p} \mathbb{1}_E$. La continuité de ϕ sur L^p donne alors

$$\lim_{k \to \infty} \nu(E_k) = \lim_{k \to \infty} \phi(\mathbb{1}_{E_k}) = \phi(\lim_{k \to \infty} \mathbb{1}_{E_k}) = \phi(\mathbb{1}_E) = \nu(E). \tag{28.480}$$

Par additivité finie de ν nous avons

$$\nu(E_k) = \sum_{i \leqslant k} \nu(A_i) \tag{28.481}$$

et en passant à la limite, $\sum_{i=1}^{\infty} \nu(A_i) = \nu(\bigcup_i A_i)$. L'application ν est donc une mesure complexe.

Mesure absolument continue En prime, si $\mu(E) = 0$ alors $\nu(E) = 0$ parce que

$$\mu(E) = 0 \Rightarrow \|\mathbb{1}_E\|_p = 0 \Rightarrow \mathbb{1}_E = 0 \text{ (dans } L^p) \Rightarrow \phi(\mathbb{1}_E) = 0 \tag{28.482}$$

47. Définition 15.198.

Utilisation de Radon-Nikodym Nous sommes donc dans un cas où $\nu \ll \mu$ et nous utilisons le théorème de Radon-Nikodym 15.199 sous la forme de la remarque 15.200 : il existe une fonction intégrable $g\colon \Omega \to \mathbb{C}$ [48] telle que pour tout $A \in \mathcal{A}$,

$$\nu(A) = \int_A \bar{g} d\mu. \tag{28.483}$$

C'est-à-dire que

$$\phi(\mathbb{1}_A) = \int_A \bar{g} d\mu = \int_\Omega \bar{g}\mathbb{1}_A d\mu. \tag{28.484}$$

Nous avons donc exprimé ϕ comme une intégrale pour les fonctions caractéristiques d'ensembles.

Pour les fonctions étagées Par linéarité si f est mesurable et étagée nous avons aussi

$$\phi(f) = \int f\bar{g} d\mu = \Phi_g(f). \tag{28.485}$$

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 : $\|f_k - f\|_\infty \to 0$ en posant $f_k = f\mathbb{1}_{|f|\leqslant k}$. Pour chaque k nous avons l'égalité

$$\Phi_g(f_k) = \phi(f_k). \tag{28.486}$$

Par ailleurs la fonction $f_k\bar{g}$ est majorée par la fonction intégrable $M\bar{g}$ et le théorème de la convergence dominée 15.184 nous donne

$$\lim_{k\to\infty} \Phi_g(f_i) = \lim_{k\to\infty} \int f_k\bar{g} = \int f\bar{g} = \Phi_g(f). \tag{28.487}$$

Et la continuité de ϕ sur L^p couplée à la convergence $f_k \xrightarrow{L^p} f$ donne $\lim_{k\to\infty} \phi(f_k) = (f)$. Bref prendre la limite dans (28.486) donne

$$\Phi_g(f) = \phi(f) \tag{28.488}$$

pour tout $f \in L^\infty(\Omega)$.

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.

De L^∞ à L^p Soit $f \in L^p$. Si nous avions une suite (f_n) dans L^∞ telle que $f_n \xrightarrow{L^p} f$ alors $\lim \phi(f_n) = \phi(f)$ par continuité de ϕ. La difficulté est de trouver une telle suite de façon à pouvoir permuter l'intégrale et la limite :

$$\lim_{n\to\infty} \int_\Omega f_n\bar{g} = \int_\Omega \lim_{n\to\infty} f_n\bar{g} = \int_\Omega f\bar{g} = \Phi_g(f). \tag{28.489}$$

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 (28.489) soit valide.

Nous allons d'abord supposer que $f \in L^p$ est positive à valeurs réelles. Nous avons alors par le théorème 15.105 qu'il existe une suite croissante de fonction étagées (et donc L^∞) telles que $f_n \to f$ ponctuellement. De plus étant donné que $|f_n| \leqslant |f|$, la proposition 28.21 nous dit que $f_n \xrightarrow{L^p} f$. Pour chaque n nous avons

$$\int_\Omega f_n\bar{g} = \phi(f_n). \tag{28.490}$$

48. On peut écrire, pour utiliser de la notation compacte que $g \in L^1(\Omega, \mathbb{C})$.

Soit g^+ la partie réelle positive de $\bar{g}$. Alors nous avons la limite croissante ponctuelle $f_n g^+ \to fg^+$ et le théorème de la convergence monotone 15.160 nous permet d'écrire

$$\lim_{n\to\infty} \int f_n g^+ = \int fg^+. \tag{28.491}$$

Faisant cela pour les trois autres parties de $\bar{g}$ nous avons prouvé que si $f \in L^p$ est réelle et positive,

$$\int f\bar{g} = \phi(f), \tag{28.492}$$

c'est-à-dire que $\Phi_g(f) = \phi(f)$.
Refaisant le tout pour les trois autres parties de f nous montrons que

$$\Phi_g(f) = \phi(f) \tag{28.493}$$

pour tout $f \in L^p(\Omega)$. Nous avons donc égalité de ϕ et Φ_g dans $(L^p)'$ et donc bijection entre $(L^p)'$ et L^q.

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

$$\|g\|_q = \|\phi\|_{(L^p)'}. \tag{28.494}$$

$\|\phi\| \leqslant \|g\|_q$ Nous savons que $\phi(f) = \int f\bar{g}$, et nous allons écrire la définition de la norme dans $(L^p)'$:

$$\|\phi\|_{(L^p)'} = \sup_{\|f\|_p=1} \big|\phi(f)\big| \tag{28.495a}$$

$$= \sup \Big| \int f\bar{g} \Big| \tag{28.495b}$$

$$\leqslant \sup \underbrace{\int |f\bar{g}|}_{=\|f\bar{g}\|_1}. \tag{28.495c}$$

Il s'agit maintenant d'utiliser l'inégalité de Hölder 28.33 :

$$\|\phi\| \leqslant \sup_{\|f\|_p=1} \|f\|_p \|\bar{g}\|_q = \|g\|_q. \tag{28.496}$$

L'inégalité dans l'autre sens sera démontrée en séparant les cas $p = 1$ et $1 < p < \infty$.

Si $p = 1$, une formule Si E est un ensemble mesurable de mesure finie, alors

$$\Big| \int_E g d\mu \Big| = \big|\phi(\mathbb{1}_E)\big|. \tag{28.497}$$

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

$$\Big| \int_\Omega \mathbb{1}_E \bar{g} d\mu \Big| = \Big| \int_E g d\mu \Big| = \big|\phi(\mathbb{1}_E)\big| \leqslant \|\phi\|_{(L^1)'} \|\mathbb{1}_E\|_1 = \|\phi\|\mu(E). \tag{28.498}$$

Nous écrivons cela dans l'autre sens :

$$\|\phi\| \geqslant \frac{1}{\mu(E)} \Big| \int_\Omega \mathbb{1}_E \bar{g} d\mu \Big| = \Big| \frac{1}{\mu(E)} \int_E \bar{g} d\mu \Big|. \tag{28.499}$$

Si nous prenons $S = \{t \in \mathbb{C} \text{ tel que } |t| \leqslant \|\phi\|\}$, c'est un fermé vérifiant que

$$\frac{1}{\mu(E)} \int_E \bar{g} d\mu \in S. \tag{28.500}$$

Voila une petite formule qui va nous aider à utiliser le lemme 28.129. Nous ne pouvons cependant pas l'utiliser immédiatement parce que l'appartenance (28.500) n'est vraie que pour les parties de mesure finie.

Si $p = 1$, conclusion[1] Pour utiliser le lemme 28.129, nous utilisons l'hypothèse que Ω est σ-fini. Soient des mesurables A_i de mesure fine tels que $\bigcup_{i\in\mathbb{N}} A_i = \Omega$.

Pour chaque i nous considérons la restriction $g_i\colon A_i \to \mathbb{C}$ de g à A_i. Par le point précédent, elle vérifie

$$\frac{1}{\mu(A_i)}\int_{A_i} \bar{g}_i d\mu = \frac{1}{\mu(A_i)}\int_{A_i} \bar{g} d\mu \in S. \tag{28.501}$$

En appliquant le lemme 28.129 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

$$\|g_i\|_\infty \leqslant \|\phi\| \tag{28.502}$$

où, dans ce contexte, $\|g_i\|_\infty$ signifie $\sup_{x\in A_i} |g_i(x)|$.

Nous avons alors

$$\|g\|_\infty = \sup_{x\in\Omega} |g(x)| = \sup_{i\in\mathbb{N}} \|g_i\|_\infty \leqslant \|\phi\|. \tag{28.503}$$

Une petite justification pour cela ? Prenons une suite x_k telle que $|g(x_k)| \to \|g\|_\infty$. Vu que les A_i recouvrent Ω, existe un naturel $i(k)$ tel que $x_k \in A_{i(k)}$. Nous avons alors

$$|g(x_k)| \leqslant \|g_{i(k)}\|_\infty \leqslant \|\phi\|. \tag{28.504}$$

Cela pour conclure que $g \in L^\infty$.

Notons que cet argument ne tient pas avec $p > 1$ parce que l'équation (28.498) terminerait sur $\|\phi\|\mu(E)^{1/p}$. Du coup l'ensemble S à prendre serait $S = \{t \in \mathbb{C} \text{ tel que } |t| \leqslant \|\phi\|\mu(E)^{1/p-1}\}$ 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.

$1 < p < \infty$ La fonction

$$\alpha(x) = \begin{cases} \frac{g(x)}{|g(x)|} & \text{si } g(x) \neq 0 \\ 1 & \text{si } g(x) = 0 \end{cases} \tag{28.505}$$

a la propriété de faire $\alpha g = |g|$ en même temps que $|\alpha(x)| = 1$ pour tout x. Nous définissons

$$E_n = \{x \text{ tel que } |g(x)| \leqslant n\} \tag{28.506}$$

et

$$f_n = \mathbb{1}_{E_n}|g^{q-1}|\alpha. \tag{28.507}$$

Ce qui est bien avec ces fonctions c'est que [49]

$$|f_n|^p = |g^{p(q-1)}|\alpha|^p = |g|^q \tag{28.508}$$

sur E_n. Dans E_n nous avons $|f_n| = |g^{q-1}| \leqslant n^{q-1}$ 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

$$\phi(f_n) = \Phi_g(f_n). \tag{28.509}$$

Par ailleurs,

$$f_n\bar{g} = \mathbb{1}_{E_n}|g^{q-1}|\frac{g}{|g|}\bar{g}, \tag{28.510}$$

49. C'est ici que nous utilisons le lien entre p et q. En l'occurrence, de $1/p + 1/q = 1$ nous déduisons $q(p-1) = p$.

donc [50]

$$\left|\int_{E_n} |g|^q d\mu\right| = |\int_\Omega f_n \bar{g} d\mu| \tag{28.511a}$$
$$= |\phi(f_n)| \tag{28.511b}$$
$$\leqslant \|\phi\|\|f_n\|_p \tag{28.511c}$$
$$= \|\phi\| \left(\int_{E_n} |f_n|^p\right)^{1/p} \tag{28.511d}$$
$$= \|\phi\| \left(\int_{E_n} |g|^q\right)^{1/p}. \tag{28.511e}$$

Nous avons de ce fait une inégalité de la forme $A \leqslant \|\phi\| A^{1/p}$ et donc aussi $A^{1/p} \leqslant \|\phi\|^{1/p} A^{1/p^2}$, et donc $A \leqslant \|\phi\|\|\phi\|^{1/p} A^{1/p^2}$. Continuant ainsi à injecter l'inégalité dans elle-même, pour tout $k \in \mathbb{N}$ nous avons :

$$\left|\int_{E_n} |g|^q d\mu\right| \leqslant \|\phi\|^{1+\frac{1}{p}+\cdots+\frac{1}{p^k}} \left(\int_{E_n} |g|^q d\mu\right)^{1/p^k}. \tag{28.512}$$

Nous pouvons passer à la limite $k \to \infty$. Sachant que $p > 1$ nous savons $A^{1/k} \to 1$ et

$$1 + \frac{1}{p} + \cdots + \frac{1}{p^k} \to \frac{p}{p-1} = q. \tag{28.513}$$

Nous avons alors

$$\int_{E_n} |g|^q d\mu \leqslant \|\phi\|^q. \tag{28.514}$$

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 $n \to \infty$ en utilisant la convergence monotone (théorème 15.160) qui donne alors $\int_\Omega |g|^q \leqslant \|\phi\|^q$ ou encore

$$\|g\|_q \leqslant \|\phi\|. \tag{28.515}$$

Ceci achève de prouver que l'application $\phi \mapsto \Phi_g$ est une isométrie, et donc le théorème.

□

Théorème 28.131.
Soit un espace mesuré $(\Omega, \mathcal{A}, \mu)$.

(1) Si $1 < p < \infty$, alors $L^p(\Omega, \mathcal{A}, \mu)$ est réflexif[51].
(2) Si $(\Omega, \mathcal{A}, \mu)$ est σ-finie, alors
(2a) $(L^1)' = L^\infty$
(2b) $L^1 \subset (L^\infty)'$.

Démonstration. En plusieurs parties, en notant toujours p et q les exposants conjugués, c'est à dire $\frac{1}{p} + \frac{1}{q} = 1$.

Pour (1) Le théorème 28.130(1) nous indique que

$$(L^p)' = L^q \tag{28.516}$$

au sens d'une bijection isométrique. 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 (28.516),

$$(L^p)'' = (L^q)' = L^p, \tag{28.517}$$

et nous avons prouvé que L^p est réflexif.

50. Dans [6], cette équation arrive sans modules, ce qui me laisse entendre que $\phi(f_n)$ est réel et positif pour pouvoir écrire que $\phi(f_n) \leqslant \|\phi\|\|f_n\|_p$, mais je ne comprends pas pourquoi.
51. Définition 12.169.

Pour (2a) Il s'agit du théorème 28.130(2).

Pour (2b)[1] Il nous reste à couvrir le cas de $(L^\infty)'$. Pour $g \in L^1$ nous prouvons que $\Phi_g \in (L^\infty)'$.

$\Phi_g(f)$ 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[52], il existe $M > 0$ tel que $|f(x)| < M$ pour tout x hors d'une partie A de mesure nulle. Nous avons alors

$$\int_\Omega |f\bar{g}| = \int_{\Omega\setminus A} |f\bar{g}| \leqslant M \int_{\Omega\setminus A} |f| = M \int_\Omega |f| < \infty. \tag{28.518}$$

Φ_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 $|f_k(x)| < \|f_k\|_{L^\infty}$ pour tout x hors de A_k. Nous avons alors

$$|\Phi_g(f_k)| \leqslant \int_{\Omega\setminus A_k} |f_k\bar{g}|d\mu \leqslant \|f_k\|_{L^\infty} \int_{\Omega\setminus A_k} |g| \leqslant \|f_k\|_{L^\infty}\|g\|_1. \tag{28.519}$$

Vu que par hypothèse $f_k \to 0$ dans L^∞, nous avons $\|f_k\|_{L^\infty \to 0}$, et donc aussi

$$|\Phi_g(f_k)| \to 0. \tag{28.520}$$

□

Proposition 28.132.
Soit $f \in L^p(\Omega)$ telle que

$$\int_\Omega f\varphi = 0 \tag{28.521}$$

pour tout $\varphi \in C_c^\infty(\Omega)$. Alors $f = 0$ presque partout.

Démonstration. Nous considérons la forme linéaire $\Phi_f \in (L^q)'$ donnée par

$$\begin{aligned} \Phi_f\colon L^p &\to \mathbb{C} \\ u &\mapsto \int_\Omega fu \end{aligned} \tag{28.522}$$

Par hypothèse cette forme est nulle sur la partie dense $C_c^\infty(\Omega)$. Si (φ_n) est une suite dans $C_c^\infty(\Omega)$ convergente vers u dans L^p, nous avons pour tout n que

$$0 = \Phi_f(\varphi_n) \tag{28.523}$$

En passant à la limite, nous voyons que Φ_f est la forme nulle. Elle est donc égale à Φ_0. La partie « unicité » du théorème de représentation de Riesz 28.130 nous indique alors que $f = 0$ dans L^p et donc $f = 0$ presque partout. □

Proposition 28.133.
Si $f \in L^1_{loc}(I)$ est telle que

$$\int_I f\varphi' = 0 \tag{28.524}$$

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

$$h = w - \psi \int_I w, \tag{28.525}$$

52. Voir les définitions 28.23 et 28.24.

qui est dans $C_c^\infty(I)$ et dont l'intégrale sur I est nulle. Par la proposition 18.2, la fonction h admet une primitive dans $C_c^\infty(I)$; et nous notons φ cette primitive. L'hypothèse appliquée à φ donne

$$0=\int_I f\varphi'=\int_I f\left(w-\psi\int_I w\right)=\int_I fw-\underbrace{\left(\int_I f(x)\psi(x)dx\right)}_{C}\left(\int_I w(y)dy\right)=\int_I w(f-C). \tag{28.526}$$

L'annulation de la dernière intégrale implique par la proposition 28.132 que $f-C=0$ dans L^2, c'est-à-dire $f=C$ presque partout. □

Dans [428], 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 18.99 dont la preuve ne me paraît pas non plus « fastidieuse mais en rien ingénieuse ».

Lemme 28.134 ([428, 1])**.**
Soient $r>0$. Il existe $\delta>0$ tel que pour tout $s,t\in\mathbb{C}$ vérifiant $|s|\leqslant 1$, $|t|\leqslant 1$ et $|s-t|\geqslant r$ nous ayons

$$\left|\frac{s+t}{2}\right|^p\leqslant(1-\delta)\frac{|s|^p+|t|^p}{2}. \tag{28.527}$$

Démonstration. Soit $r>0$. La partie de $\mathbb{C}^2$ donnée par

$$D=\{(s,t)\in\mathbb{C}^2 \text{ tel que } |s|\leqslant 1,|t|\leqslant 1,|s-t|\geqslant r\} \tag{28.528}$$

est compacte. En effet elles est bornée (par la sphère de rayon $\sqrt{2}$) et fermée comme intersection de fermée[53]. Nous considérons la fonction $\Delta\colon D\to\mathbb{R}$ donnée par

$$\left|\frac{s+t}{2}\right|^p=\Delta(s,t)\frac{|s|^p+|t|^p}{2}. \tag{28.529}$$

Si vous voulez une expression explicite,

$$\Delta(s,t)=\frac{2^{p-1}|s+t|^p}{|s|^p+|t|^p}. \tag{28.530}$$

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 18.99 nous dit que la fonction $z\mapsto|z|^p$ est strictement convexe. En prenant la définition 18.92 de la stricte convexité avec $\theta=\frac{1}{2}$, nous trouvons que

$$\Delta(s,t)<1 \tag{28.531}$$

pour tout $(s,t)\in D$. Vu que par ailleurs Δ est une fonction continue sur le compact D, elle atteint un minimum dans D. Soit Δ_0 ce minimum quivérifie forcément $\Delta_0<1$.

En posant $1-\delta=\Delta_0$ nous avons le résutat. □

28.12 Théorèmes de Hahn-Banach

Théorème 28.135 (Hahn-Banach[429, 117])**.**
Soit E, un espace vectoriel réel et une application $p\colon E\to\mathbb{R}$ satisfaisant

(1) $p(\lambda x)=\lambda p(x)$ pour tout $x\in E$ et pour tout $\lambda>0$,

(2) $p(x+y)\leqslant p(x)+p(y)$ pour tout $x,y\in E$.

53. Lemme 7.3 suivit du théorème de Borel-Lebesgue 8.9.

Soit de plus $G \subset E$ un sous-espace vectoriel muni d'une application $g\colon G \to \mathbb{R}$ vérifiant $g(x) \leqslant p(x)$ pour tout $x \in G$. Alors il existe $f \in \mathcal{L}(E, \mathbb{R})$ telle que $f(x) = g(x)$ pour tout $x \in G$ et $f(x) \leqslant p(x)$ pour tout $x \in E$.

Démonstration. Si h une application linéaire définie sur un sous-espace de E, nous notons D_h ledit sous-espace.

Un ensemble inductif Nous considérons P, l'ensemble des fonctions linéaires suivant

$$P = \left\{ h\colon D_h \to \mathbb{R} \text{ tel que } \begin{cases} G \subset D_h \\ h(x) = g(x) \quad \forall x \in G \\ h(x) \leqslant p(x) \quad \forall x \in D_h \end{cases} \right\} \tag{28.532}$$

Cet ensemble est non vide parce que g est dedans. Nous le munissons de la relation d'ordre $h_1 \leqslant h_2$ 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

$$\begin{aligned} h\colon D_h &\to \mathbb{R} \\ x &\mapsto l(x) \quad \text{si } x \in D_l \end{aligned} \tag{28.533}$$

Cela est bien définit parce que si $x \in D_l \cap D_{l'}$ alors, vu que Q est totalement ordonné (i.e. $l \leqslant l'$ ou $l' \leqslant l$), 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 $h \geqslant l$ pour tout $l \in Q$. Cela montre que P est inductif (définition 1.14). Le lemme de Zorn 1.15 nous dit alors que P possède un maximum f qui va être la réponse à notre théorème.

Le support de f La fonction f est dans P ; donc $f(x) \leqslant p(x)$ pour tout $x \in D_h$ et $f(x) = g(x)$ pour tout $x \in G$. 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

$$h(x + tx_0) = f(x) + t\alpha \tag{28.534}$$

où α est une constante que nous allons fixer plus tard.

Nous commençons par prouver que f est dans P. Nous devons prouver que

$$h(x + tx_0) = f(x) + t\alpha \leqslant p(x + tx_0) \tag{28.535}$$

Pour cela nous allons commencer par fixer α pour avoir les relations suivantes :

$$\begin{cases} f(x) + \alpha \leqslant p(x + x_0) & \text{(28.536a)} \\ f(x) - \alpha \leqslant p(x - x_0) & \text{(28.536b)} \end{cases}$$

pour tout $x \in D_f$. Ces relations sont équivalentes à demander α tel que

$$\begin{cases} \alpha \leqslant p(x + x_0) - f(x) & \text{(28.537a)} \\ \alpha \geqslant f(x) - p(x - x_0) & \text{(28.537b)} \end{cases}$$

Nous nous demandons donc s'il existe un α qui satisfasse

$$\sup_{y \in D_f} \big(f(y) - p(y - x_0)\big) \leqslant \alpha \leqslant \inf_{z \in D_f} \big(p(z + x_0) - f(z)\big). \tag{28.538}$$

Ou encore nous devons prouver que pour tout $y, z \in D_f$,

$$p(z + x_0) - f(x) \geqslant f(y) - p(y - x_0) \geqslant 0. \tag{28.539}$$

Par les propriétés de p et de f,

$$p(z+x_0)+p(y-x_0)-f(z)-f(y)\geqslant p(z+y)-f(z+y)\geqslant 0. \tag{28.540}$$

La dernière inégalité est le fait que $f\in P$. Un choix de α donnant les inéquations (28.536) est donc possible.

À partir des inéquations (28.536) nous obtenons la relation (28.535) de la façon suivante. Si $t>0$ nous multiplions l'équation (28.536a) par t :

$$tf(x)+t\alpha\leqslant tp(x+x_0). \tag{28.541}$$

Et nous écrivons cette relation avec x/t au lieu de x en tenant compte de la linéarité de f :

$$f(x)+t\alpha\leqslant tp\Big(\frac{x}{t}+x_0\Big)=p(x+tx_0). \tag{28.542}$$

Avec $t<0$, c'est similaire, en faisant attention au sens des inégalités.

Nous avons donc construit $h\colon D_h\to\mathbb{R}$ 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.

□

Définition 28.136.
Un espace topologique est ***localement convexe*** *si tout point possède un système fondamental de voisinages formé de convexes.*

Définition 28.137 (Hyperplan qui sépare)**.**
Soit E un espace vectoriel topologique ainsi que A, B des sous-ensembles de E. Nous disons que l'hyperplan d'équation $f=\alpha$ ***sépare au sens large*** *les parties A et B si $f(x)\leqslant\alpha$ pour tout $x\in A$ et $f(x)\geqslant\alpha$ pour tout $x\in B$.*

La séparation est ***au sens strict*** *s'il existe $\epsilon>0$ tel que*

$$f(x)\leqslant\alpha-\epsilon \qquad \text{pour tout } x\in A \tag{28.543a}$$
$$f(x)\geqslant\alpha+\epsilon \qquad \text{pour tout } x\in B. \tag{28.543b}$$

Théorème 28.138 (Hahn-Banach, première forme géométrique[117])**.**
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.

Théorème 28.139 (Hahn-Banach, seconde forme géométrique)**.**
Soient un espace vectoriel topologique localement convexe[54] *ainsi que des convexes non vides disjoints A et B tels que A soit compact et B soit fermé. Alors il existe un hyperplan fermé qui sépare strictement A et B.*

Démonstration. Vu que B est fermé, A est dans l'ouvert $E\backslash B$. Donc si $a\in A$, il existe un voisinage ouvert convexe de a inclus dans A. Soit U_a un voisinage ouvert et convexe de 0 tel que $(a+U_a)\cap B=\varnothing$.

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 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

$$A\subset\bigcup_{i=1}^{n}(a_i+V_{a_i}). \tag{28.544}$$

54. Définition 28.136.

Nous posons alors

$$V = \bigcap_{i=1}^{n} V_{a_i}. \tag{28.545}$$

Cet ensemble est non vide parce et 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

$$x \in a_i + U_{a_i} + V \subset a_i + V_{a_i} + V_{a_i} \subset a_i + U_{a_i} \subset E\backslash B. \tag{28.546}$$

Donc $(A + V) \cap B = \varnothing$. L'ensemble $A + V$ est alors un ouvert convexe disjoint de B. Par la première forme géométrique du théorème de Hahn-Banach 28.138 nous avons un hyperplan qui sépare $A + V$ de B au sens large : il existe $f \in E'\backslash\{0\}$ tel que $f(a) + f(v) \leqslant f(b)$ 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 E. □

28.13 Théorème de Tietze

Définition 28.140.
*Si E et F sont des espaces normés, une application $f : E \to F$ est **presque surjective** s'il existe $\alpha \in]0, 1[$ et $C > 0$ tels que pour tout $y \in \overline{B_F(0, 1)}$, il existe $x \in \overline{B_E(0, C)}$ tel que $\|y - f(x)\| \leqslant \alpha$.*

Lemme 28.141 ([43])**.**
Soient E et F des espaces de Banach et $f \in L(E, F)$[55]*. Si f est presque surjective, alors*

(1) f est surjective

(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 $x \in B\big(0, \frac{C}{1-\alpha}\big)$ 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 $\|y - f(x_1)\| \leqslant \alpha$. Ensuite nous considérons la récurrence

$$x_n \in \overline{B_E(0, C)} \tag{28.547}$$

tel que

$$\Big\|y - \sum_{i=1}^{n} \alpha^{i-1} f(x_i)\Big\| \leqslant \alpha^n \tag{28.548}$$

Pour montrer que cela existe nous supposons que la série est déjà construire jusqu'à $n > 1$:

$$\frac{1}{\alpha^n}\Big(y - \sum_{i=1}^{n} \alpha^{i-1} f(x_i)\Big) \in \overline{B_F(0, 1)} \tag{28.549}$$

À partir de là, par presque surjectivité il existe un $x_{n+1} \in \overline{B_E(0, C)}$ tel que

$$\Big\|\frac{y - \sum_{i=1}^{n} \alpha^{i-1} f(x_i)}{\alpha^n} - f(x_{n+1})\Big\| \leqslant \alpha. \tag{28.550}$$

En multipliant par α^n, le terme $\alpha^n f(x_{n+1})$ s'intègre bien dans la somme :

$$\Big\|y = \sum_{i=1}^{n+1} \alpha^{i-1} f(x_i)\Big\| \leqslant \alpha^{n+1}. \tag{28.551}$$

55. L'ensemble des applications linéaires continues

Nous nous intéressons à une éventuelle limite à la somme des $\alpha^{n-1}x_n$. D'abord nous avons la majoration $\|\alpha^{n-1}x_n\| \leqslant \alpha^{n-1}C$, et vu que par la définition de la presque surjectivité $0 < \alpha < 1$, la série

$$\sum_{n=1}^{\infty} \alpha^{n-1}x_n \tag{28.552}$$

converge absolument [56] parce que la suite des normes est une suite géométrique de raison α. 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

$$x = \sum_{n=1}^{\infty} \alpha^{n-1}x_n \in E, \tag{28.553}$$

et en termes de normes, ça vérifie

$$\|x\| \leqslant \sum_{n=1}^{\infty} \alpha^{n-1}\|x_n\| \leqslant C \sum_{n=1}^{\infty} \alpha^{n-1} = \frac{C}{1-\alpha}. \tag{28.554}$$

Donc c'est bon pour avoir $x \in B\big(0, \frac{C}{1-\alpha}\big)$. Nous devons encore vérifier que $y = f(x)$. Pour cela nous remarquons que

$$\|y - f\Big(\sum_{n=1}^{N} \alpha^{n-1}x_n\Big)\| \leqslant \alpha^N. \tag{28.555}$$

Nous pouvons prendre la limite $N \to \infty$ et permuter f avec la limite (par continuité de f). Vu que $0 < \alpha < 1$ nous avons

$$\|y - f(x)\| = 0. \tag{28.556}$$

□

Théorème 28.142 (Tietze[43, 276]).
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

$$\begin{aligned} T\colon (C_b^0(X, \mathbb{R}), \|.\|_\infty) &\to (C_b^0(Y, \mathbb{R}), \|.\|_\infty) \\ f &\mapsto f|_Y. \end{aligned} \tag{28.557}$$

L'application T est évidemment linéaire. Elle est de plus borné pour la norme opérateur usuelle donnée par la proposition 12.10 parce que $\|T(f)\| \leqslant \|f\| < \infty$. L'application T est alors continue par la proposition 12.25.

Presque surjection Soit $g \in C_b^0(Y, \mathbb{R})$ avec $\|g\|_\infty \leqslant 1$. Nous posons

$$Y^+ = \{x \in Y \text{ tel que } \frac{1}{3} \leqslant g(x) \leqslant 1\} \tag{28.558a}$$

$$Y^- = \{x \in Y \text{ tel que } -1 \leqslant g(x) \leqslant -\frac{1}{3}\}. \tag{28.558b}$$

Nous considérons alors

$$\begin{aligned} f\colon X &\to \mathbb{R} \\ x &\mapsto \frac{1}{3}\frac{d(x, Y^-) - d(x, Y^+)}{d(x, Y^-) + d(x, Y^+)} \end{aligned} \tag{28.559}$$

Vu qu'en valeur absolue le dénominateur est plus grand que le numérateur nous avons $\|f\|_\infty \leqslant \frac{1}{3}$. Notons que

— Si $x \in Y^+$ alors $f(x) = \frac{1}{3}$ et $g(x) \in [\frac{1}{3}, 1]$;

56. Définition 12.59.

— 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[57] $g(x) \in [-\frac{1}{3}, \frac{1}{3}]$ et donc $\big|f(x) - g(x)\big| \leqslant \big|f(x)\big| + \big|g(x)\big| \leqslant \frac{2}{3}$.

Dans les deux cas nous avons $\big|f(x) - g(x)\big| \in [0, \frac{2}{3}]$ pour tout $x \in X$. Cela prouve que

$$\|T(f) - g\|_{Y,\infty} \leqslant \frac{2}{3}. \tag{28.560}$$

En résumé nous avons pris g dans la boule $\overline{B(0,1)}$ de $\big(C_b^0(Y,\mathbb{R}), \|.\|_\infty\big)$ et nous avons construit une fonction f dans la boule $\overline{B(0,\frac{1}{3})}$ de $\big(C_b^0(X,\mathbb{R}), \|.\|_\infty\big)$ telle que $\|T(f) - g\|_\infty \leqslant \frac{2}{3}$. L'application T est donc une presque surjection avec $\alpha = \frac{1}{3}$ et $C = \frac{2}{3}$.

Prolongement dans les boules unité fermées La proposition 13.282 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 28.141 nous dit alors que T est surjective et que pour tout $g \in \overline{B(0,1)}$, il existe

$$f \in \overline{B\left(0, \frac{1/3}{1 - \frac{2}{3}}\right)} = \overline{B(0,1)}. \tag{28.561}$$

telle que $g = T(f)$.

Prolongement pour les boules ouvertes Jusqu'à présent nous avons montré qu'une fonction $g \in \overline{B(0,1)}$ admet une prolongement continu dans $\overline{B(0,1)}$. Nous allons montrer que si g est dans la boule ouverte $B(0,1)$ de $\big(C_b^0(Y,\mathbb{R}), \|.\|_\infty\big)$ alors g admet un prolongement dans la boule ouverte $B(0,1)$ de $\big(C_b^0(X,\mathbb{R}), \|.\|_\infty\big)$.

Soit $g \in B_{C_b^0(Y)}(0,1)$ et son prolongement $h \in \overline{B_{C_b^0(X)}(0,1)}$. Si $\|h\|_\infty < 1$ alors le résultat est vrai. Sinon nous considérons l'ensemble

$$Z = \{x \in X \text{ tel que } |h(x)| = 1\}. \tag{28.562}$$

Nous avons $Y \cap Z = \varnothing$ parce que nous avons $h = g$ sur Y et nous avons choisi $\|g\|_\infty < \infty$. Par ailleurs Y est fermé par hypothèse et Z est fermé parce que h est continue ; par conséquent $Y \cap Z$ est fermé, donc[58]

$$\bar{Y} \cap \bar{Z} = Y \cap Z = \varnothing. \tag{28.563}$$

Nous posons

$$\begin{aligned} u\colon X &\to \mathbb{R}^+ \\ x &\mapsto \frac{d(x,Z)}{d(x,Y) + d(x,Z)} \end{aligned} \tag{28.564}$$

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 $|f(x)| < |h(x)| \leqslant 1$. Donc $\|f\|_\infty < 1$.

Nous avons donc trouvé qu'une fonction dans la boule ouverte $B_{C_b^0(Y)}(0,1)$ se prolonge en une fonction dans la boule ouverte $B_{C_b^0(X)}(0,1)$.

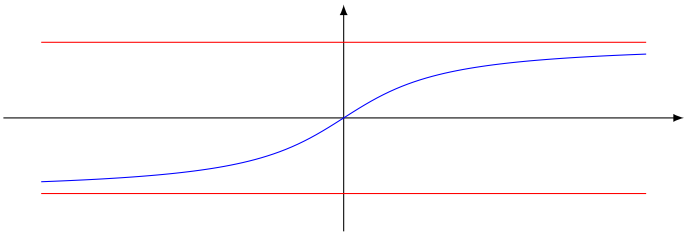
Le cas non borné Soit enfin $g_0 \in C^0(Y,\mathbb{R})$. Nous allons nous ramener au cas de la boule unité ouverte en utilisant un homéomorphisme $\phi\colon \mathbb{R} \to]-1,1[$. L'application $g = \phi \circ g_0$ est dans la boule unité ouvert de $C^0(Y,\mathbb{R})$ et donc admet un prolongement f dans la boule unité ouverte de $C^0(X)$. L'application $f_0 = \phi^{-1} \circ f$ est un prolongement continu de g_0.

□

57. Nous rappelons que $\|g\| = 1$, donc $g(x)$ est forcément ente -1 et 1.

58. Si vous avez l'intention de dire que $\overline{Y \cap Z} = \bar{Y} \cap \bar{Z} = Y \cap Z = \varnothing$, allez d'abord voir l'exemple 8.59. Ici c'est correct parce que Y et Z sont fermés.

Un homéomorphisme $\phi\colon \mathbb{R} \to]-1,1[$ est donné par exemple par la fonction $\phi(t) = \frac{2}{\pi}\arctan(t)$ dont le graphique est donné ci-dessous :



28.14 Espace de Schwartz

Pour un multiindice $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$, nous notons

$$\partial^\alpha \varphi = \partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d} \varphi \tag{28.565}$$

pour peu que la fonction φ soit $|\alpha| = \alpha_1 + \cdots + \alpha_d$ fois dérivable.

Définition 28.143.
*Soit $\Omega \subset \mathbb{R}^d$. L'**espace de Schwartz** $\mathscr{S}(\Omega)$ est le sous-ensemble de $C^\infty(\Omega)$ des fonctions dont toutes les dérivées décroissent plus vite que tout polynôme :*

$$\mathscr{S}(\Omega) = \big\{\varphi \in C^\infty(\Omega) \text{ tel que } \forall \alpha, \beta \in \mathbb{N}^d, p_{\alpha,\beta}(\varphi) < \infty\big\} \tag{28.566}$$

où nous avons considéré

$$p_{\alpha,\beta}(\varphi) = \sup_{x\in\Omega} |x^\beta(\partial^\alpha\varphi)(x)| = \|x^\beta \partial^\alpha \varphi\|_\infty. \tag{28.567}$$

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 28.144
La fonction e^{-x^2} est une fonction à décroissance rapide sur $\mathbb{R}$. △

Définition 28.145.
*Une fonction $f\colon \mathbb{R}^d \to \mathbb{C}$ 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 $\|x\| \geqslant r$.*

Proposition 28.146.
Une fonction Schwartz est à décroissance rapide.

Démonstration. Nous commençons par considérer un polynôme P donné par

$$P(x) = \sum_k c_k x^{\beta_k} \tag{28.568}$$

où les β_k sont des multiindices, les c_k sont des constantes et la somme est finie. Nous avons la majoration

$$\sup_{x\in\mathbb{R}^d} |\varphi(x)P(x)| \leqslant \sum_k \sup_x \left|c_k\varphi(x)x^{\beta_k}\right| \leqslant \sum_k |c_k| p_{0,\beta_k}(\varphi) < \infty. \tag{28.569}$$

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 ayons $|\varphi(x)P(x)| \leqslant M_P$ et donc

$$|\varphi(x)| \leqslant \frac{M_P}{|P(x)|} = \frac{1}{|\frac{1}{M_P}P(x)|}. \tag{28.570}$$

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 $|P(x)| \geqslant C|Q(x)|$ pour tout $|x| \geqslant r_C$. En particulier pour $|x| \geqslant r_{M_P}$ nous avons

$$|P(x)| \geqslant M_P|Q(x)| \tag{28.571}$$

et donc, pour ces x,

$$|\varphi(x)| \leqslant \frac{1}{|\frac{1}{M_P}P(x)|} \leqslant \frac{1}{|Q(x)|}. \tag{28.572}$$

La première inégalité est valable pour tout x, et la seconde pour $\|x\| \geqslant r_{M_P}$. □

Corollaire 28.147 ([1]).
Soit φ une fonction Schwartz sur $\mathbb{R}^m \times \mathbb{R}^n$. Alors la fonction

$$y \mapsto \sup_{x\in\mathbb{R}^n} |\varphi(x, y)| \tag{28.573}$$

est intégrable.

Démonstration. Soit un polynôme Q en la variable y. Par la proposition 28.146, il existe $r > 0$ tel que

$$|\varphi(x, t)| < \frac{1}{Q(y)} \tag{28.574}$$

pour tout $\|(x, y)\| > r$. A fortiori l'inégalité tient pour tout $|y| > r$. Donc

$$\int_{\mathbb{R}^m} \sup_{x\in\mathbb{R}^n} |\varphi(x, y)|dy = \int_{\|y\|\leqslant r} \sup_x |\varphi(x, y)|dy + \int_{\|y\|>r} \sup_x |\varphi(x, y)|dy. \tag{28.575}$$

La première intégrale est bornée par $\text{Vol}\big(B(0, r)\big)\|f\|_\infty$ 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. □

28.14.1 Topologie

Lemme-définition 28.148.
Les $p_{\alpha,\beta}$ donnés par l'équation (28.567) ci-dessus sont des semi-normes[59]*. La topologie considérée sur $\mathscr{S}(\mathbb{R}^d)$ est celle des semi-normes $p_{\alpha,\beta}$.*

28.149.
Nous avons un enchaînement de résultats qui nous aident à prouver la continuité d'une application $T : \mathscr{S}(\mathbb{R}^d) \to X$.

(1) La topologie de $\mathscr{S}(\mathbb{R}^d)$ est donnée par une famille dénombrable de semi-normes. Donc la proposition 9.79 nous dit que $\mathscr{S}(\mathbb{R}^d)$ est métrisable.

(2) La proposition 9.49 nous dit alors que si X est métrique, toute application séquentiellement continue $T : \mathscr{S}(\mathbb{R}^d) \to X$ est continue.

(3) Donc si X est métrique, il suffit de prouver que pour $f_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} 0$ nous avons $T(f_n) \xrightarrow{X} 0$ où $f_n : \mathscr{S}(\mathbb{R}^d) \to X$. Dans les cas usuels, T sera une distribution et $X = \mathbb{C}$.

(4) En vertu de la proposition 9.77, la convergence $f_n \xrightarrow{\mathscr{S}(\mathbb{R}^d)} 0$ signifie que pour tout choix de multiindice α et β, $p_{\alpha,\beta}(f_n) \to 0$, c'est-à-dire

$$\|x^\beta \partial^\alpha f_n\|_\infty \to 0. \tag{28.576}$$

59. Définition 9.74.

(5) Et enfin, la technique pour montrer que $T\colon \mathscr{S}(\mathbb{R}^d) \to \mathbb{C}$ est continue est de montrer que sous l'hypothèse d'avoir (28.576) pour tout choix de α et β, nous avons $T(f_n) \to 0$ dans $\mathbb{C}$.

Lemme 28.150 ([4]).
La topologie sur $\mathscr{S}(\mathbb{R}^d)$ est donnée aussi par les semi-normes

$$q_{n,m} = \max_{|\alpha|\leqslant n} \sup_{x\in\mathbb{R}^d} \big(1+\|x\|\big)^m \big|\partial^\alpha \varphi(x)\big|. \tag{28.577}$$

Autrement dit, une suite $\varphi_n \overset{\mathscr{S}(\mathbb{R}^d)}{\to} 0$ si et seulement si $q_{n,m(\varphi)} \to 0$ pour tout n et m.

Le fait que les $q_{n,m}(\varphi)$ restent bornés est la proposition 28.146. Cependant ce lemme est plus précis parce qu'en disant seulement que φ est majoré par des polynôme, nous ne disons pas que les polynômes correspondants aux φ_n tendent vers zéro si $\varphi_n \overset{\mathscr{S}}{\to} 0$. Et d'ailleurs on ne sait pas très bien ce que signifierait $P_n \to 0$ pour une suite de polynômes.

Proposition 28.151.
Pour $p \in [1,\infty]$, l'espace $\mathscr{S}(\mathbb{R}^d)$ 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 $f_n \overset{\mathscr{S}}{\to} 0$, alors $f_n \overset{L^p}{\to} 0$.

Commençons par $p = \infty$. Alors $\|f_n\|_\infty = p_{0,0}(f_n) \to 0$ parce que si $f_n \overset{\mathscr{S}}{\to} 0$, alors en particulier $p_{0,0}(f_n) \to 0$.

Au tour de $p < \infty$ maintenant. Nous savons qu'en dimension d, la fonction

$$x \mapsto \frac{1}{(1+\|x\|)^s} \tag{28.578}$$

est intégrable dès que $s > d$. Pour toute valeur de m nous avons

$$\|\varphi\|_p^p = \int_{\mathbb{R}^d} |\varphi(x)|^p dx = \int_{\mathbb{R}^d} \frac{\big|(1+\|x\|)^m \varphi(x)\big|^p}{\big(1+\|x\|\big)^{mp}} \leqslant \int_{\mathbb{R}^d} \frac{q_{0,m}(\varphi)^p}{\big(1+\|x\|\big)^{mp}}. \tag{28.579}$$

En choisissant m de telle sorte que $mp > d$, nous avons convergence de l'intégrale et donc $\|\varphi\|_p < \infty$. Nous retenons que

$$\|\varphi\|_p^p \leqslant C q_{0,m}(\varphi)^p \tag{28.580}$$

pour une certaine constance C et un bon choix de m.

Ceci prouve que $\mathscr{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$. Nous devons encore vérifier que l'inclusion est continue. Si $\varphi_n \overset{\mathscr{S}}{\to} 0$, alors en particulier nous avons $q_{0,m}(\varphi_n) \to 0$ par le lemme 28.150. Par conséquent la majoration (28.580) nous dit que $\|\varphi_n\|_p \to 0$ également.

□

En résumé, si $\varphi_n \overset{\mathscr{S}(\mathbb{R}^d)}{\to} \varphi$ alors $\varphi_n \overset{L^p}{\to} \varphi$.

Théorème 28.152 ([263]).
Soit μ une mesure sur les boréliens de $\mathbb{R}^n$ finie sur les compacts. Alors $\mathscr{D}(\mathbb{R}^n)$ est dense dans $L^1(\mathbb{R}^n, \mathcal{B}or(\mathbb{R}^n), \mu)$.

Proposition 28.153 ([430]).
La partie $\mathscr{D}(\mathbb{R}^d)$ est dense dans $\mathscr{S}(\mathbb{R}^d)$.

Démonstration. Soit $f \in \mathscr{S}(\mathbb{R}^d)$, et ϕ, une fonction de $\mathscr{D}(\mathbb{R}^d)$ telle que $\phi(x) = 1$ pour $|x| \leqslant 1$ (l'existence de telles fonctions est discutée en 16.13.1). Soit aussi $\phi_k(x) = \phi(x/k)$. Nous posons

$$f_k(x) = \phi_k(x) f(x), \tag{28.581}$$

et nous allons prouver que pour tout multiindices α et γ,

$$p_{\alpha,\gamma}(f_k - f) = \|x^\gamma \partial^\alpha (f_k - f)\|_\infty \to 0. \tag{28.582}$$

Pour cela nous allons noter $\beta \leqslant \alpha$ lorsque β est un multiindice contenu dans α. En utilisant la dérivée du produit nous avons

$$\begin{aligned}
(\partial^\alpha f_k)(x) &= \sum_{\beta\leqslant\alpha} (\partial^{\alpha-\beta}\phi_k)(x)\partial^\beta f(x) && (28.583a)\\
&= \sum_{\beta\leqslant\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x) && (28.583b)\\
&= \sum_{\beta<\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x) + \phi(x/k)(\partial^\alpha f)(x). && (28.583c)
\end{aligned}$$

Nous devons donc étudier et majorer

$$\begin{aligned}
\sup_{x\in\mathbb{R}^d} |x^\gamma \partial^\alpha (f_k - f)| \leqslant \sup \Big| x^\gamma \sum_{\beta<\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x)\Big| \\
+ \sup \big| x^\gamma \big(\phi(x/k) - 1\big)(\partial^\alpha f)(x)\big|
\end{aligned} \tag{28.584}$$

En ce qui concerne le second terme, soit $\epsilon > 0$, vu que f est Schwartz, il existe R tel que

$$|x^\gamma(\partial^\alpha f)(x)| < \epsilon \tag{28.585}$$

dès que $\|x\| > R$. En prenant $k > R$,

$$|x^\gamma(\partial^\alpha f)(x)| \begin{cases} = 0 & \text{si } \|x\| < R \\ \leqslant \epsilon & \text{si } \|x\| > R. \end{cases} \tag{28.586}$$

En ce qui concerne le premier terme,

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^d} \Big| x^\gamma \sum_{\beta<\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x)\Big| && (28.587a)\\
&\qquad \leqslant \frac{1}{k} \sup_x \Big| \sum_{\beta<\alpha} (\partial^{\alpha-\beta}\phi)(x/k)(x^\gamma \partial^\beta f)(x)\Big| && (28.587b)\\
&\qquad = \frac{1}{k} \sup_x \Big| \sum_{\beta<\alpha} (\partial^{\alpha-\beta}\phi)(x/k) p_{\beta,\gamma}(f)\Big| && (28.587c)
\end{aligned}$$

La somme ne contient qu'un nombre fini de β différents, donc nous pouvons considérer un nombre K qui majore tous les $p_{\beta,\gamma}(f)$ en même temps. La partie avec ϕ 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 β. Vu que la somme est finie, elle est majorée par ce nombre multiplié par le nombre de termes dans la somme et au final

$$\sup_{x\in\mathbb{R}^d} \Big| x^\gamma \sum_{\beta<\alpha} k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\phi)(x/k)(\partial^\beta f)(x)\Big| \leqslant \frac{K'}{k}. \tag{28.588}$$

La limite $k \to \infty$ ne fait alors plus de doutes. □

Remarque 28.154.

Vu la topologie de $\mathscr{S}(\mathbb{R}^d)$ (définition 28.148), la convergence $f_k \xrightarrow{\mathscr{S}(\mathbb{R}^d)} f$ peut être exprimée par le fait que pour tout k, l,

$$t^k f_n^{(l)} \xrightarrow{unif} t^k f^{(l)}. \tag{28.589}$$

C'est-à-dire convergence uniforme de toutes les dérivées multipliées par n'importe quel polynôme.

28.14.2 Produit de convolution

Proposition 28.155 (Stabilité de Schwartz par convolution[60] [417]).
Si $\varphi \in L^1(\mathbb{R}^d)$ et $\psi \in \mathscr{S}(\mathbb{R}^d)$, alors $\varphi * \psi \in \mathscr{S}(\mathbb{R}^d)$.

Démonstration. Nous devons prouver que

$$p_{\alpha,\beta}(\varphi * \psi) = \sup_{x\in\mathbb{R}^d} |x^\beta(\partial^\alpha(\varphi * \psi))(x)| \tag{28.590}$$

est borné pour tout multiindices α et β. En appliquant $|\alpha|$ fois la proposition 28.57, nous mettons toutes les dérivées sur ψ : $\partial^\alpha(\varphi * \psi) = (\varphi * \partial^\alpha\psi)$. Cela étant fait, nous majorons

$$\begin{aligned}
\big|x^\beta(\varphi * \partial^\alpha\psi)(x)\big| &\leqslant |x^\beta| \int_{\mathbb{R}^d} |\varphi(y)| \underbrace{\big|(\partial^\alpha\psi)(x-y)\big|}_{\leqslant\|\partial^\alpha\psi\|_\infty} dy\Big| && (28.591a)\\
&\leqslant |x^\beta|\|\partial^\alpha\psi\|_\infty \int_{\mathbb{R}^d} |\varphi(y)|dy && (28.591b)\\
&\leqslant p_{\alpha,\beta}(\psi)\|\varphi\|_{L^1}. && (28.591c)
\end{aligned}$$

Par conséquent, $p_{\alpha,\beta}(\varphi * \psi) \leqslant \|\varphi\|_{L^1} p_{\alpha,\beta}(\psi) < \infty$. □

28.15 Théorème de Montel

Théorème 28.156 (Montel[43]).
Soient Ω un ouvert de $\mathbb{C}$ et $\mathcal{F}$ une famille de fonctions holomorphes sur Ω, uniformément bornée sur tout compact de Ω. Alors de toute suite dans $\mathcal{F}$ nous pouvons extraire une sous-suite convergeant uniformément sur tout compact de Ω.

Démonstration. **Un ensemble équicontinu** Nous commençons par prendre une suite de compacts dans Ω comme dans le lemme 9.59, et une suite δ_n de réels strictement positifs tels que

$$B(z, 2\delta_n) \subset K_{n+1} \tag{28.592}$$

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 27.36 nous donne

$$f(x) - f(y) = \frac{1}{2\pi i}\int_{\partial B}\left(\frac{f(\xi)}{\xi - x} - \frac{f(\xi)}{\xi - y}\right)d\xi = \frac{x-y}{2\pi i}\int_{\partial B}\frac{f(\xi)}{(\xi-x)(\xi-y)}d\xi \tag{28.593}$$

Nous majorons ça par

$$|f(x) - f(y)| \leqslant \frac{|x-y|}{2\pi}\int_{\partial B}\frac{|f(\xi)|}{2\delta_n^2}d\xi \leqslant \frac{|x-y|}{\delta_n}M_n. \tag{28.594}$$

Justifications :

— $|\xi - x| = 2\delta_n$ et $|\xi - y| \geqslant \delta_n$ parce que ξ est au mieux sur le rayon passant par x et y.
— $|f(\xi)| \leqslant M_n$ 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

$$\mathcal{F}_n = \{f|_{K_n} \text{ tel que } f \in \mathcal{F}\} \tag{28.595}$$

est équicontinu. Il est aussi équiborné par hypothèse.

60. Définition 28.52.

Application du théorème d'Ascoli L'ensemble $\mathcal{F}_n$ vérifie les hypothèses du théorème d'Ascoli 28.5. 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 $\bar{\mathcal{F}}_n$, c'est-à-dire uniformément. Autrement dit il existe une fonction strictement croissante $\varphi\colon \mathbb{N} \to \mathbb{N}$ telle que la suite $k \mapsto f_{\varphi(k)}$ converge uniformément sur K_n. La limite n'est cependant pas spécialement dans $\mathcal{F}_n$.

L'argument diagonal La suite $k \mapsto f_{\varphi_1 \circ \ldots \varphi_k(k)}$ converge uniformément sur tous les K_n. Si K est un compact de Ω, alors les petites propriétés sympas du lemme 9.59 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 Ω. □

Corollaire 28.157 ([43]).
Soient Ω un ouvert connexe borné de $\mathbb{C}$ et $a \in \Omega$. Soit f holomorphe sur Ω 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 Ω vers la fonction constante a.

Démonstration. Nous considérons un voisinage de a inclus dans Ω ; sachant que $|f(a)| < 1$, nous trouvons un voisinage encore plus petit de a sur lequel $|f'(z)| < 1$. Soit donc r tel que $\overline{B(a,r)} \subset \Omega$ et tel que $|f'(z)| < 1$ sur $\overline{B(a,r)}$. Étant donné que $f'(z)$ est continue sur le compact $\overline{B(a,r)}$, nous en prenons le maximum λ (qui est strictement inférieur à 1) et nous avons au final

$$|f'(z)| \leqslant \lambda < 1 \tag{28.596}$$

pour tout $z \in \overline{B(a,r)}$. Le théorème des accroissements finis 13.252 nous dit que

$$\big|f(z) - a\big| \leqslant \lambda |z - a| \tag{28.597}$$

pour tout $z \in \overline{B(a,r)}$. C'est ici que nous utilisons l'hypothèse de convexité de Ω. Nous montrons alors par récurrence que

$$\big|f^n(z) - a\big| \leqslant \lambda^n |z - a| \leqslant \lambda^n r \leqslant r. \tag{28.598}$$

L'ensemble $A = \{f^n \text{ tel que } n \geqslant 1\}$ 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

$$f^n(z) \in \overline{B(a,r)}. \tag{28.599}$$

La suite (f^n) est donc uniformément bornée sur tout compact de $B(a,r)$. Le théorème de Montel 28.156 nous indique que l'on peut extraire une sous-suite convergente uniformément sur tout compact. Au vu de (28.598) cette convergence ne peut avoir lieu que vers une fonction g qui vaut la constante a sur $B(a,r)$.

D'autre par la fonction g est holomorphe en tant que limite uniforme de fonctions holomorphes, théorème 27.69. Or une fonction holomorphe constante sur un ouvert est constante sur tout son domaine d'holomorphie (principe d'extension analytique, théorème 18.126). □

28.16 Espaces de Bergman

Source : [276].

Soit Ω un borné dans $\mathbb{C}$ et D le disque unité ouvert de $\mathbb{C}$.

Définition 28.158.
*L'**espace de Bergman** sur Ω, noté $A^2(\Omega)$ est l'espace des fonctions holomorphes sur Ω qui sont en même temps dans $L^2(\Omega)$.*

Nous mettons sur $A^2(\Omega)$ le produit scalaire usuel hérité de L^2 :

$$\langle f, g\rangle = \int_\Omega f(z)\overline{g(z)}dz. \tag{28.600}$$

Lemme 28.159.
Soient un compact $K \subset \Omega$ et une fonction $f \in A^2(\Omega)$. Alors

$$\max_{z\in K} |f(z)| \leqslant \frac{1}{\sqrt{\pi}} \frac{1}{d(K, \partial\Omega)} \|f\|_2. \tag{28.601}$$

Démonstration. Soient $a \in \Omega$ et $r > 0$ tels que $B(a, r) \subset \Omega$. Nous considérons aussi $\rho \leqslant r$. La formule de Cauchy (27.36) nous donne

$$f(a) = \frac{1}{2\pi i} \int_{B(a,\rho)} \frac{f(\xi)}{\xi - a} f\xi = \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho e^{i\theta})d\theta \tag{28.602}$$

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

$$\frac{r^2}{2} f(a) = \int_0^r f(a)\rho d\rho, \tag{28.603}$$

et d'y substituer la valeur de $f(a)$ que nous venons de calculer :

$$\begin{aligned}
\frac{r^2}{2} f(a) &= \int_0^r \frac{1}{2\pi} \int_0^{2\pi} f(a + \rho e^{i\theta})d\theta\rho d\rho && (28.604a)\\
&= \frac{1}{2\pi} \int_{B(a,r)} f(z)dz && \text{passage aux polaires} && (28.604b)\\
&= \frac{1}{2\pi} \langle 1, f\rangle_B && \text{produit scalaire sur } B(a, r) && (28.604c)\\
&\leqslant \frac{1}{2\pi} \sqrt{\langle 1, 1\rangle_B \langle f, f\rangle_B} && (28.604d)
\end{aligned}$$

Nous avons donc

$$r^2 f(a) \leqslant \frac{1}{\pi} \sqrt{\langle 1, 1\rangle_B \langle f, f\rangle_B}, \tag{28.605}$$

et donc

$$\pi r^2 f(a) \leqslant \sqrt{\pi r^2} \|f\|_2, \tag{28.606}$$

parce que $\langle f, f\rangle_B \leqslant \|f\|_2^2$. En effet le produit scalaire $\|.\|_2$ est donné par une intégrale sur Ω alors que $B(a, r) \subset \Omega$ et que la fonction qu'on y intègre est positive (c'est $|f(z)|^2$). En simplifiant,

$$f(a) \leqslant \frac{1}{\sqrt{\pi} r} \|f\|_2. \tag{28.607}$$

Mais r a été choisi pour avoir $B(a, r) \subset \Omega$, donc $r \leqslant d(a, \partial\Omega)$ et

$$|f(a)| \leqslant \frac{1}{d(a, \partial\Omega)\sqrt{\pi}} \|f\|_2. \tag{28.608}$$

Maintenant si nous prenons $a \in K$, nous avons encore la minoration $d(a, \partial K) \leqslant d(a, \partial\Omega)$ et donc

$$|f(a)| \leqslant \frac{1}{d(a, \partial K)\sqrt{\pi}} \|f\|_2. \tag{28.609}$$

□

Théorème 28.160.
Soit Ω 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

$$e_n(z) = \sqrt{\frac{n+1}{\pi}} z^n \tag{28.610}$$

pour $n \geqslant 0$.

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 28.159 que

$$\max_{z\in K} \big|f_n(z) - f_m(z)\big| \leqslant \frac{1}{\sqrt{\pi} d(K, \partial\Omega)} \|f_n - f_m\|_2. \tag{28.611}$$

Donc f_n converge uniformément sur K. Par le théorème de Weierstrass 27.69, la fonction f est holomorphe. Il existe donc une fonction holomorphe f qui est limite uniforme sur tout compact de Ω 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 28.19 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 $(e_n)_{n\geqslant 0}$ est une base orthonormale. En ce qui concerne les produits scalaires,

$$\begin{aligned}
\langle e_m, e_n\rangle &= \sqrt{\frac{(m+1)(n+1)}{\pi}} \int_D z^n \overline{z^m} dz && (28.612a)\\
&= \sqrt{\frac{(m+1)(n+1)}{\pi^2}} \int_0^1 \rho\, d\rho \int_0^{2\pi} d\theta \rho^{m+n} e^{i\theta(n-m)} && (28.612b)\\
&= \sqrt{\frac{(m+1)(n+1)}{\pi^2}} \frac{1}{m+n+2} \underbrace{\int_0^{2\pi} e^{i\theta(n-m)} d\theta}_{2\pi\delta_{mn}} && (28.612c)\\
&= \sqrt{\frac{(n+1)^2}{\pi^2}} \frac{1}{2n+2} 2\pi\delta_{nm} && (28.612d)\\
&= \delta_{nm}. && (28.612e)
\end{aligned}$$

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

$$\|f\|_2^2 = \sum_{n=0}^{\infty} |\langle f, e_n\rangle|^2, \tag{28.613}$$

parce que le point (5) du théorème 26.44 nous indique que ce sera suffisant pour avoir une base hilbertienne.

Étant donné que f est holomorphe sur D, le théorème 27.13 nous développe f en série entière :

$$f(z) = \sum_{k=0}^{\infty} a_k z^k. \tag{28.614}$$

En permutant la somme avec le produit scalaire,

$$c_n(f) = \int_D f(z)\bar{e}_n(z) = \sqrt{\frac{n+1}{\pi}} \int_D f(z)\bar{z}^n dz. \tag{28.615}$$

Afin de profiter de la convergence uniforme de la série (28.614) 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

$$g_k(z)=\begin{cases}f(z)\bar{z}^n & \text{si } |z|<1-1/k\\ 0 & \text{sinon.}\end{cases}\tag{28.616}$$

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 $g_k(z)\to f(z)\bar{z}^n$. Le théorème de la convergence dominée permet alors de permuter l'intégrale et la limite $k\to\infty$. Cela nous permet d'écrire

$$c_n(f)=\sqrt{\frac{n+1}{\pi}}\lim_{r\to1^-}\int_{|z|<r}\bar{z}^n f(z)dz=\sqrt{\frac{n+1}{\pi}}\lim_{r\to1^-}\int_{|z|<r}\sum_{k=0}^{\infty}a_k z^k\bar{z}^n.\tag{28.617}$$

Par la convergence uniforme de la série entière *à l'intérieur* du disque D nous pouvons permuter l'intégrale et la somme (proposition 16.30) :

$$c_n(f)=\sqrt{\frac{n+1}{\pi}}\lim_{r\to1^-}\sum_{k=0}^{\infty}a_k\int_{|z|<r}z^k\bar{z}^n dz.\tag{28.618}$$

L'intégrale proprement dite est vite calculée et vaut

$$\int_{|z|<1}\bar{z}^n z^k dz=\frac{\pi r^{2n+2}}{n+1}\delta_{kn}.\tag{28.619}$$

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 :

$$c_n(f)=\sqrt{\frac{n+1}{\pi}}\lim_{r\to1^-}\sum_k a_k\frac{\pi r^{2n+2}}{n+1}\delta_{kn}=\sqrt{\frac{\pi}{n+1}}a_n.\tag{28.620}$$

Nous effectuons le même genre de calculs pour évaluer $\|f\|_2^2$:

$$\begin{aligned}\|f\|_2^2&=\int_D|f(z)|^2dz && && \text{(28.621a)}\\ &=\lim_{r\to1^-}\int_{|z|<r}f(z)\sum_{k=0}^{\infty}\bar{a}_k\bar{z}_k dz && && \text{(28.621b)}\\ &=\lim_{r\to1^-}\sum_{k=0}^{\infty}\bar{a}_k\int_{|z|<r}f(z)\bar{z}^k dz && \text{permuter } \textstyle\sum \text{ et } \int && \text{(28.621c)}\\ &=\lim_{r\to1^-}\sum_{k=0}^{\infty}\bar{a}_k a_k\frac{\pi r^{2k+2}}{k+1} && \text{intégrale déjà faite.} && \text{(28.621d)}\end{aligned}$$

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

$$\|f\|_2^2=\lim_{r\to1^-}\sum_{k=0}^{\infty}|c_k(f)|^2r^{2k+2}.\tag{28.622}$$

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 26.24 nous indique que $\sum_j|c_k(f)|^2\leqslant\|f\|_2^2$. Le théorème de la convergence dominée nous permet d'inverser la limite et la somme pour obtenir le résultat attendu :

$$\|f\|_2^2=\sum_{k=0}^{\infty}|c_k(f)|^2.\tag{28.623}$$

□