Le Frido 2017,
volume 1
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 2018. 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/mazhe.pdf`

3. Et bien entendu les sources LaTeX sur github :

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

Copyright 2011-2017 Laurent Claessens, Carlotta Donadello, Lilian Besson
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”.

ISBN : 979-10-97085-01-8

Thèmes

Ceci est une sorte d'index thématique.

1 : intégration
2 : suites et séries
3 : polynôme de Taylor
4 : séries de Fourier
5 : normes
6 : topologie produit
7 : espaces métriques, normés
8 : norme opérateur
9 : gaussienne
10 : produit de compact
11 : densité
12 : espaces de fonctions
13 : fonctions Lipschitz
14 : formule des accroissements finis
15 : différentiabilité
16 : points fixes
17 : théorème de Stokes, Green et compagnie
18 : permuter des limites
19 : applications continues et bornées
20 : inégalités
21 : connexité
22 : suite de Cauchy, espace complet
23 : application réciproque
24 : tribu, algèbre de parties, λ-systèmes et co.
25 : déduire la nullité d'une fonction depuis son intégrale
26 : mesure et intégrale
27 : équations différentielles
28 : injections
29 : logarithme
30 : inversion locale, fonction implicite
31 : convexité
32 : dualité
33 : opérations sur les distributions
34 : transformée de Fourier
35 : convolution
36 : méthode de Newton
37 : méthodes de calcul
38 : définie positive
39 : norme matricielle et rayon spectral
40 : série de matrices
41 : rang
42 : extension de corps et polynômes
43 : décomposition de matrices
44 : systèmes d'équations linéaires
45 : racines de polynôme et factorisation de polynômes
46 : formes bilinéaires et quadratiques
47 : arithmétique modulo, théorème de Bézout
48 : polynômes
49 : invariants de similitude
50 : diagonalisation
51 : endomorphismes cycliques
52 : déterminant
53 : polynôme d'endomorphismes
54 : exponentielle de matrice
55 : sous-groupes
56 : groupe symétrique
57 : action de groupe
58 : classification de groupes
59 : produit semi-direct de groupes
60 : théorie des représentations
61 : isométries
62 : caractérisation de distributions en probabilités
63 : théorème central limite
64 : lemme de transfert
65 : probabilités et espérances conditionnelles
66 : dénombrements
67 : enveloppes
68 : équations diophantiennes

Thème 1 : intégration

(1) L'existence d'une primitive pour toute fonction continue est le théorème 18.10.

(2) La définition d'une primitive est la définition 12.135.

(3) Primitive et intégrale, proposition 13.204.

(4) Intégrale impropre, définition 13.217.

Thème 2 : suites et séries De façon un peu contre-intuitive, les suites et séries sont surtout traitées dans le chapitre sur les espaces vectoriels. Section 10.5 pour les suites et 10.6 pour les séries.

(1) La définition de la somme d'une infinité de termes est donnée par la définition 10.43.

(2) La définition de la convergence absolue est la définition 10.45.

(3) Convergence d'une série implique convergence vers zéro du terme général, proposition 10.53.
(4) Quelque séries usuelles dans $\mathbb{R}$ dans la section 10.7.3.

Pour les sommes infinies l'associativité et la commutativité dans une série sont perdues. Néanmoins, il subsiste que

(1) si la série converge, on peut regrouper ses termes sans modifier la convergence ni la somme (associativité) ;
(2) si la série converge absolument, on peut modifier l'ordre des termes sans modifier la convergence ni la somme (commutativité, proposition 9.176).

Une somme indexée par un ensemble quelconque est la définition 9.178.

Thème 3 : polynôme de Taylor Énoncés :

(1) Énoncé : théorème 12.263.
(2) De classe C^2 sur $\mathbb{R}^n$, proposition 12.267.
(3) Avec un reste donné par un point dans $]x, a[$, proposition 12.271.
(4) Le polynôme de Taylor généralise à l'utilisation de toutes les dérivées disponibles le résultat de développement limité donné par la proposition 12.120.
(5) Pour les fonctions holomorphes, il y a le théorème 25.27 qui donne une série de Taylor sur un disque de convergence.

Utilisation :

(1) Il est utilisé pour justifier la méthode de Newton autour de l'équation (32.92).
(2) On utilise pas mal de Taylor dans les résultats liant extrema et différentielle/Hessienne. Par exemple la proposition 18.68.

Thème 4 : séries de Fourier

— Formule sommatoire de Poisson, proposition 27.9.
— Inégalité isopérimétrique, théorème 26.21.
— Fonction continue et périodique dont la série de Fourier ne converge pas, proposition 26.18.
— Nous allons montrer la convergence de $\sum_{k\in\mathbb{Z}} c_k(f)e^{inx}$ vers $f(x)$ dans divers cas :
  (1) Si f est continue et périodique, convergence au sens de Cesaro, théorème de Fejèr 26.5.
  (2) Convergence au sens $L^2\Big([0, 2\pi]\Big)$ dans le théorème 24.57.
  (3) Si f est continue, périodique et sa série de Fourier converge uniformément, théorème 26.13.
  (4) Si f est périodique et la série des coefficients converge absolument pour tout x, proposition 26.14.
  (5) Si f est périodique et de classe C^1, théorème 26.15.

  Il est cependant faux de croire que la continuité et la périodicité suffisent à obtenir une convergence, comme le montre dans la proposition 26.18.

Thème 5 : normes

Définition définition 6.31.

Équivalence de norme (1) Définition de l'équivalence de norme 10.1.
(2) La proposition 10.2 sur l'équivalence des normes dans $\mathbb{R}^n$.
(3) Toutes les normes sur un espace vectoriel de dimension finie sont équivalentes, théorème 10.3.
(4) Montrer que le problème $a - b$ est stable dans l'exemple 32.26.
(5) La proposition 10.15 donnant $\rho(A) \leqslant \|A\|$ utilise l'équivalence de toutes les normes sur un espace vectoriel de dimension finie.

Norme opérateur voir le thème 8.

Thème 6 : topologie produit

(1) La définition de la topologie produit est 6.5.

(2) Pour les espaces vectoriels normés, le produit est donné par la définition 10.73.

(3) L'équivalence entre la topologie de la norme produit et la topologie produit est le lemme 10.75.

Thème 7 : espaces métriques, normés

(1) Le théorème-définition 6.19 donne la topologie sur un espace métrique.

(2) La définition de la convergence d'une suite est la définition 6.4.

(3) Dans un espace vectoriel normé, une application est continue si et seulement si elle est bornée, proposition 10.18.

Thème 8 : norme opérateur Pour la norme matricielle et le rayon spectral, voir le thème 39.

(1) Définition 10.5.

(2) Définition d'une norme d'algèbre 10.9.

(3) Pour des espaces vectoriels normée, être borné est équivalent à être continu : proposition 10.18.

(4) Le lemme à propos d'exponentielle de matrice 15.94 donne :

$$\|e^{tA}\| \leqslant P\big(|t|\big) \sum_{i=1}^{r} e^{t\,\mathrm{Re}(\lambda_i)}. \tag{0.1}$$

Thème 9 : gaussienne

(1) Le calcul de l'intégrale

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi} \tag{0.2}$$

est fait de deux façons dans l'exemple 13.233. Les deux utilisent le théorème de Fubini 13.230.

(2) Le lemme 27.15 calcule la transformée de Fourier de $g_\epsilon(x) = e^{-\epsilon\|x\|^2}$ qui donne $\hat{g}_\epsilon(\xi) = \big(\frac{\pi}{\epsilon}\big)^{d/2} e^{-\|\xi\|^2/4\epsilon}$.

(3) Le lemme 27.18 donne une suite régularisante à base de gaussienne.

(4) Elle est utilisée pour régulariser une intégrale dans la preuve de la formule d'inversion de Fourier 27.20

Thème 10 : produit de compact

(1) Les produits d'espaces métriques compacts sont compacts ; c'est le théorème de Tykhonov. Nous verrons ce résultat dans les cas suivants.
   — $\mathbb{R}$, lemme 6.122.
   — Produit fini d'espaces métriques compacts, théorème 6.162.
   — Produit dénombrable d'espaces métrique compacts, théorème 6.164.

Thème 11 : densité

(1) Densité des polynômes dans $C^0\big([0,1]\big)$, théorème de Bernstein 34.133.

(2) Densité de $\mathscr{D}(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$ pour $1 \leqslant p < \infty$, théorème 24.33.

(3) Densité de $\mathscr{S}(\mathbb{R}^d)$ dans l'espace de Sobolev $H^s(\mathbb{R}^d)$, proposition 29.17.

(4) Densité de $\mathscr{D}(\mathbb{R}^d)$ dans l'espace de Sobolev $H^s(\mathbb{R}^d)$, proposition 29.19.
Cela est utilisé pour le théorème de trace 29.21.

(5) Les applications étagées dans les applications mesurables (qui plus est avec limite croissante), théorème fondamental d'approximation 24.36.
(6) Les fonctions continues à support compact dans $L^2(I)$, théorème 24.37.
(7) Les polynômes trigonométriques sont denses dans $L^p(S^1)$ pour $1 \leqslant p < \infty$. Deux démonstrations indépendantes par le théorème 26.6 et le théorème 24.51.

Les densités sont bien entendu utilisées pour prouver des formules sur un espace en sachant qu'elles sont vraies sur une partie dense. Mais également pour étendre une application définie seulement sur une partie dense. C'est par exemple ce qui est fait pour définir la trace γ_0 sur les espaces de Sobolev $H^s(\mathbb{R}^d)$ en utilisant le théorème d'extension 18.112.

Comme presque tous les théorèmes importants, le théorème de Stone-Weierstrass possède de nombreuses formulations à divers degrés de généralité.

— Le lemme 18.1 le donne pour la racine carré.
— Le théorème 18.6 donne la densité des polynômes dans les fonctions continues sur un compact.
— Le théorème 18.4 est une généralisation qui donne la densité uniforme d'une sous-algèbre de $C(X, \mathbb{R})$ dès que X sépare les points.
— Le lemme 26.1 est une version pour les polynômes trigonométriques.
— Le lemme 18.1 est une cas particulier du théorème 18.6, mais nous en donnons une démonstration indépendante afin d'isoler la preuve de la généralisation 18.4. Une version pour les polynômes trigonométriques sera donnée dans le lemme 26.1.

Thème 12 : espaces de fonctions En ce qui concerne les densités, voir le thème 11.

Topologie Les espaces de fonctions sont souvent munis de topologies définies par des semi-normes.

(1) La topologie des semi-normes est la définition 6.190.
(2) La définition 28.3 donne les topologies sur $C^\infty(\Omega)$, $\mathscr{D}(K)$ et $\mathscr{D}(\Omega)$.
(3) La topologie $*$-faible sur $\mathscr{D}'(\Omega)$ est donnée par la définition 28.10.

L'espace $L^2\big([0, 2\pi]\big)$ C'est un espace très important, entre autres parce qu'il est de Hilbert et est bien adapté à la transformée de Fourier.

(1) Définition de $L^2(\Omega, \mu)$, 24.53.
(2) Le produit scalaire $\langle f, g\rangle$ est donné en (24.187) et la base trigonométrique est (24.188).
(3) La densité des polynômes trigonométriques dans $L^p(S^2)$ est le théorème 24.51 ou le théorème 26.6, au choix.
(4) Une conséquence de cette densité est que le système trigonométrique est une base hilbertienne de L^2 par le lemme 24.56.

L'espace L^2 est discuté en analyse fonctionnelle, en 24.5 parce que l'étude de L^2 utilise entre autres l'inégalité de Hölder 24.22.

Le fait que L^2 soit une espace de Hilbert est utilisé dans la preuve du théorème de représentation de Riesz 24.60.

Thème 13 : fonctions Lipschitz

(1) Définition : 12.243.
(2) La notion de Lipschitz est utilisée pour définir la stabilité d'un problème, définition 32.25.

Thème 14 : formule des accroissements finis Il en existe plusieurs formes :

(1) Une version adaptée aux espaces de dimension finie est le théorème 12.241.
(2) Pour les fonctions $\mathbb{R} \to \mathbb{R}$ en le théorème 12.129.
(3) Une généralisation pour les intervalles non bornés : théorème 12.130.
(4) Espaces vectoriels normés, théorème 12.229

Thème 15 : différentiabilité

(1) La recherche d'extrema d'une fonction sur $\mathbb{R}^n$ passe par la seconde différentielle, proposition 18.68.

(2) Lien entre différentielle seconde (Hessienne) et convexité en la proposition 18.90 et le corollaire 18.92.

(3) La différentielle est liée aux dérivées partielles par les formules données au lemme 12.178

$$df_a(u) = \frac{\partial f}{\partial u}(a) = \frac{d}{dt}\Big[f(a+tu)\Big]_{t=0} = \sum_{i=1}^{m} u_i \frac{\partial f}{\partial x_i}(a) = \nabla f(a) \cdot u. \tag{0.3}$$

Thème 16 : points fixes

(1) Il y a plusieurs théorèmes de points fixes.

Théorème de Picard 18.23 donne un point fixe comme limite d'itérés d'une fonction Lipschitz. Il aura pour conséquence le théorème de Cauchy-Lipschitz 18.37, l'équation de Fredholm, théorème 18.28 et le théorème d'inversion locale dans le cas des espaces de Banach 18.47.

Théorème de Brouwer qui donne un point fixe pour une application d'une boule vers elle-même. Nous allons donner plusieurs versions et preuves.

(a) Dans $\mathbb{R}^n$ en version C^∞ via le théorème de Stokes, proposition 18.31.

(b) Dans $\mathbb{R}^n$ en version continue, en s'appuyant sur le cas C^∞ et en faisant un passage à la limite, théorème 18.32.

(c) Dans $\mathbb{R}^2$ via l'homotopie, théorème 25.19. Oui, c'est très loin. Et c'est normal parce que ça va utiliser la formule de l'indice qui est de l'analyse complexe [1].

Théorème de Markov-Kakutani 18.35 qui donne un point fixe à une application continue d'un convexe fermé borné dans lui-même. Ce théorème donnera la mesure de Haar 18.36 sur les groupes compacts.

Théorème de Schauder 18.33 qui est une version valable en dimension infinie du théorème de Brouwer.

(2) Pour les équations différentielles

(a) Le théorème de Schauder a pour conséquence le théorème de Cauchy-Arzela 18.43 pour les équations différentielles.

(b) Le théorème de Schauder 18.33 permet de démontrer une version du théorème de Cauchy-Lipschitz (théorème 18.37) sans la condition Lipschitz, mais alors sans unicité de la solution. Notons que de ce point de vue nous sommes dans la même situation que la différence entre le théorème de Brouwer et celui de Picard : hors hypothèse de type «contraction», point d'unicité.

(3) En calcul numérique
- La convergence d'une méthode de point fixe est donnée par la proposition 32.47.
- La convergence quadratique de la méthode de Newton est donnée par le théorème 32.53.
- En calcul numérique, section 32.5
- Méthode de Newton comme méthode de point fixe, sous-section 32.6.2.

(4) D'autres utilisation de points fixes.
- Processus de Galton-Watson, théorème 36.47.
- Dans le théorème de Max-Milgram 23.56, le théorème de Picard est utilisé.

Thème 17 : théorème de Stokes, Green et compagnie

(1) Forme générale, théorème 14.48.

(2) Rotationnel et circulation, théorème 22.8.

Le théorème de Stokes peut être utilisé pour montrer le théorème de Brower, proposition 18.31.

1. On aime bien parce que ça ne demande pas Stokes, mais quand même hein, c'est pas gratos non plus.

Thème 18 : permuter des limites

(1) Les théorèmes sur les fonctions définies par des intégrales, section 14.10. Nous avons entre autres
   (a) $\partial_i \int_B f = \int_B \partial_i f$, avec B compact, proposition 14.72.
   (b) Si f est majorée par une fonction ne dépendant pas de x, nous avons le théorème 14.62.
   (c) Si l'intégrale est uniformément convergente, nous avons le théorème 14.63.
   (d) Pour dériver $\int_B g(t,z)dt$ avec B compact dans $\mathbb{R}$ et $g\colon \mathbb{R}\times\mathbb{C}\to\mathbb{C}$, il faut aller voir la proposition 25.17.

(2) Théorème de la convergence monotone, théorème 13.129.

(3) Le théorème de Fubini permet non seulement de permuter des intégrales, mais également des sommes parce que ces dernières peuvent être vues comme des intégrales sur $\mathbb{N}$ muni de la tribu des parties et de la mesure de comptage. Nous utilisons cette technique pour permute une somme et une intégrale dans l'équation (25.122).
   — le théorème de Fubini-Tonelli 13.227 demande que la fonction soit mesurable et positive ;
   — le théorème de Fubini 13.230 demande que la fonction soit intégrable (mais pas spécialement positive) ;
   — le corollaire 13.229 demande l'intégrabilité de la valeur absolue des intégrales partielles pour déduire que la fonction elle-même est intégrable.

Thème 19 : applications continues et bornées

(1) Une application linéaire non continue : exemple 10.20 de $e_k \mapsto ke_k$. Les dérivées partielles sont calculées en (23.131).

(2) Une autre application linéaire non continue en l'exemple 10.21 : la dérivation sur les polynômes.

(3) Une application linéaire est bornée si et seulement si elle est continue, proposition 10.18.

Thème 20 : inégalités

Inégalité de Jensen (1) Une version discrète pour $f\big(\sum_i \lambda_i x_i\big)$, la proposition 18.93.

(2) Une version intégrale pour $f\big(\int \alpha d\mu\big)$, la proposition 24.20.

(3) Une version pour l'espérance conditionnelle, la proposition 34.52.

Inégalité de Minkowsky (1) Pour une forme quadratique q sur $\mathbb{R}^n$ nous avons $\sqrt{q(x+y)} \leqslant \sqrt{q(x)} + \sqrt{q(y)}$. Proposition 9.254.

(2) Si $1 \leqslant p < \infty$ et si $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ alors $\|f+g\|_p \leqslant \|f\|_p + \|g\|_p$. Proposition 24.26.

(3) L'inégalité de Minkowsky sous forme intégrale s'écrit sous forme déballée

$$\left[\int_X \Big(\int_Y |f(x,y)|d\nu(y)\Big)^p d\mu(x)\right]^{1/p} \leqslant \int_Y \Big(\int_X |f(x,y)|^p d\mu(x)\Big)^{1/p} d\nu(y). \tag{0.4}$$

ou sous forme compacte

$$\Big\| x \mapsto \int_Y f(x,y)d\nu(y)\Big\|_p \leqslant \int_Y \|f_y\|_p d\nu(y) \tag{0.5}$$

Transformée de Fourier Pour tout $f \in L^1(\mathbb{R}^n)$ nous avons $\|\hat{f}\|_\infty \leqslant \|f\|_1$, lemme 27.6.

Inégalité des normes Inégalité de normes : si $f \in L^p$ et $g \in L^1$, alors $\|f * g\|_p \leqslant \|f\|_p \|g\|_1$, proposition 24.40.

Thème 21 : connexité

(1) Définition 6.84
(2) Le groupe $\mathrm{SL}(n,\mathbb{K})$ est connexe par arcs : proposition 12.333.
(3) Le groupe $\mathrm{GL}(n,\mathbb{C})$ est connexe par arcs : proposition 12.334.
(4) Le groupe $\mathrm{GL}(n,\mathbb{C})$ est connexe par arcs, proposition 12.334.
(5) Le groupe $\mathrm{GL}(n,\mathbb{R})$ a exactement deux composantes connexes par arcs, proposition 12.335.
(6) Le groupe $\mathrm{O}(n,\mathbb{R})$ n'est pas connexe, lemme 12.329.
(7) Les groupe $\mathrm{U}(n)$ et $\mathrm{SU}(n)$ sont connexes par arcs, lemme 12.330.
(8) Le groupe $\mathrm{SO}(n)$ est connexe mais ce n'est pas encore démontré, proposition 12.331.
(9) Connexité des formes quadratiques de signature donnée, proposition 18.101.

Thème 22 : suite de Cauchy, espace complet Nous parlons d'espaces topologiques complets. À ne pas confondre avec un espace mesuré complet, définition 13.50.

(1) La définition 6.39 donne la notion de suite de Cauchy dans un espace métrique.
(2) La définition 6.37 donne la notion de suite de τ-Cauchy dans un espace vectoriel topologique.
(3) Deux espaces métriques (avec une distance) peuvent être isomorphes en tant qu'espaces topologiques, mais ne pas avoir les mêmes suites de Cauchy, exemple 6.42.
(4) La proposition 6.52 donne l'équivalence entre les suites de Cauchy et les suites τ-Cauchy dans le cas des espaces vectoriels topologiques *normés*.
(5) L'exemple 6.42 est un exemple pire que simplement une suite de Cauchy qui ne converge pas. Le problème de convergence de cette suite n'est pas simplement que la limite n'est pas dans l'espace ; c'est que la suite de Cauchy donnée ne convergerait même pas dans $\mathbb{R}$.
(6) Le théorème 18.119 est un théorème de complétion d'un espace métrique.

Quelque espaces qui sont complets sont listés ci-dessous. Attention : la complétude est bien une propriété de la norme ; le même ensemble peut être complet pout une norme et pas pour une autre. Si on vous demande «est-ce que tel espace est complet ?» vous devez répondre en demandant «pour quelle topologie ?» ... sauf si la topologie est évidente dans le contexte.

(1) Les réels $\mathbb{R}$, théorème 6.57.
(2) La proposition 12.299 donne quelque espaces complets. Soit X un espace topologique métrique (Y,d) un espace espace métrique complet. Alors les espaces
   (a) $\left(C_b^0(X,Y),\|.\|_\infty\right)$
   (b) $\left(C_0^0(X,Y),\|.\|_\infty\right)$
   (c) $\left(C_0^k(X,Y),\|.\|_\infty\right)$
   sont complets.
(3) Le lemme 12.300 dit que $\left(C^0(A,B),\|.\|_\infty\right)$ est complet dès que A est compact et B est complet.
(4) L'espace $\mathscr{D}(K)$ est complet tant pour la topologie des semi-normes que pour la topologie métrique (qui sont les mêmes). C'est la proposition 28.7.
(5) L'espace $\mathscr{S}(\Omega)$ est complet et métrisable par la proposition 28.48.

La limite uniforme d'une suite de fonctions dérivables n'est pas spécialement dérivable. Même si les fonctions sont de classe C^∞, la limite n'est pas spécialement mieux que continue. En effet, le théorème de Stone-Weierstrass 18.6 nous dit que les polynômes (qui sont C^∞) sont denses dans les fonctions continues sur un compact pour la norme uniforme. Vous ne pouvez donc pas espérer que $\left(C^p(X,Y),\|.\|_\infty\right)$ soit complet en général.

Thème 23 : application réciproque

(1) Définition 6.187.
(2) Dans le cas des réels, des exemples sont donnés en 12.7.
(3) Continuité, proposition 6.189.
(4) Théorème de la bijection 12.41 (qui contient aussi de la continuité).
(5) Dérivabilité, proposition 12.42.

Thème 24 : tribu, algèbre de parties, λ-systèmes et co. Il existe des centaines de notions de mesures et de classes de parties.

(1) Le plus souvent lorsque nous parlons de mesure est que nous parlons de mesure positive, définition 13.23 sur un espace mesuré avec une tribu, définition 13.1.
(2) Une mesure extérieure est la définition 13.17
(3) Une algèbre de partie : définition 13.18. Une mesure sur une algèbre de parties : définition 13.20. L'intérêt est que si on connait une mesure sur une algèbre de parties, elle se prolonge en une mesure sur la tribu engendrée par le théorème de prolongement de Hahn 13.60.
(4) Un λ-système : définition 13.37.
(5) Une mesure complexe : définition 13.147.

Thème 25 : déduire la nullité d'une fonction depuis son intégrale Des résultats qui disent que si $\int f = 0$ c'est que $f = 0$ dans un sens ou dans un autre.

(1) Il y a le lemme 13.123 qui dit ça.
(2) Un lemme du genre dans L^2 existe aussi pour $\int f\varphi = 0$ pour tout φ. C'est le lemme 24.43.
(3) Et encore un pour L^p dans la proposition 24.65.
(4) Si $\int f\chi = 0$ pour tout χ à support compact alors $f = 0$ presque partout, proposition 24.1.
(5) La proposition 24.17 donne $f = 0$ dans L^p lorsque $\int fg = 0$ pour tout $g \in L^q$ lorsque l'espace est σ-fini.
(6) Une fonction $h \in C_c^\infty(I)$ admet une primitive dans $C_c^\infty(I)$ si et seulement si $\int_I h = 0$. Théorème 18.11.

Thème 26 : mesure et intégrale

(1) Mesure de Lebesgue, définition 13.172
(2) Intégrale associée à une mesure, définition 13.113
(3) Mesure à densité, définition 13.137.

Thème 27 : équations différentielles L'utilisation des théorèmes de point fixe pour l'existence de solutions à des équations différentielles est fait dans le chapitre sur les points fixes.

(1) Le théorème de Schauder a pour conséquence le théorème de Cauchy-Arzela 18.43 pour les équations différentielles.
(2) Le théorème de Schauder 18.33 permet de démontrer une version du théorème de Cauchy-Lipschitz (théorème 18.37) sans la condition Lipschitz
(3) Le théorème de Cauchy-Lipschitz 18.37 est utilisé à plusieurs endroits :
— Pour calculer la transformée de Fourier de $e^{-x^2/2}$ dans le lemme 27.15.
(4) Théorème de stabilité de Lyapunov 30.34.
(5) Le système proie prédateurs, Lokta-Voltera 30.35
(6) Équation de Schrödinger, théorème 30.41.
(7) L'équation $(x - x_0)^\alpha u = 0$ pour $u \in \mathscr{D}'(\mathbb{R})$, théorème 28.56.
(8) La proposition 30.37 donne un résultat sur $y'' + qy = 0$ à partir d'une hypothèse de croissance.
(9) Équation de Hill $y'' + qy = 0$, proposition 30.39.

Thème 28 : injections

(1) L'espace de Sobolev $H^1(I)$ s'injecte de façon compacte dans $C^0(\bar{I})$, proposition 29.8.
(2) L'espace de Sobolev $H^1(I)$ s'injecte de façon continue dans $L^2(I)$, proposition 29.8.
(3) L'espace $L^2(\Omega)$ s'injecte continument dans $\mathscr{D}'(\Omega)$ (les distributions), proposition 28.13.

Thème 29 : logarithme

(1) L'exemple 15.63 donne une série pour $\ln(1-x)$.

(2) La proposition 9.196 dit que toute matrice complexe admet un logarithme. En particulier une série explicite est donné pour le logarithme d'un bloc de Jordan.

(3) Le logarithme pour les réels strictement positifs est donné en la définition 15.57.

(4) Sur les complexes, le logarithme $\ln\colon \mathbb{C}^* \to \mathbb{C}$ est la définition 25.54. Attention : ce n'est pas la seule définition possible.

(5) La série harmonique diverge à vitesse logarithmique, et la série des inverses des nombres premiers, c'est encore plus lent : théorème 15.68.

Thème 30 : inversion locale, fonction implicite

Des énoncés (1) Inversion locale dans $\mathbb{R}^n$: théorème 18.46. Pour un Banach c'est le théorème 18.47.

(2) Fonction implicite dans un Banach : théorème 18.48.

Des utilisations (1) Utilisé pour montrer que le flot d'une équation différentielle est un C^p-difféomorphisme local, voir 30.30.

(2) Pour le théorème de Von Neumann 18.60.

Thème 31 : convexité L'essentiel des résultats sur les fonctions convexes sont dans la sections 18.9.

(1) Définition des fonctions convexes : 18.74 et 18.88 en dimension supérieure.

(2) En termes de différentielles, 18.89 pour la différentielle première et 18.92 pour la Hessienne.

(3) Une courbe paramétrée convexe est la définition 19.84.

(4) L'enveloppe convexe d'une courbe fermée simple et convexe : 19.86.

(5) Courbure et convexité d'une courbe paramétrée : section 19.12.4.

(6) Une courbe paramétrée convexe est localement le graphe d'une fonction convexe, lemme 19.85.

(7) La convexité est utilisée dans la méthode du gradient à pas optimal de la proposition 18.98.

Thème 32 : dualité Ne pas confondre dual algébrique et dual topologique d'un espace vectoriel.

Dual topologique et algébrique Ils sont définis par 7.40. Le dual algébrique est l'ensemble des formes linéaires, et le dual topologique ne considère que les formes linéaires continues (en dimension infinie, les applications linéaires ne sont pas toutes continues).

Topologie Une topologie possible sur le dual d'un espace vectoriel topologique est celle $*$-faible de la définition 6.196.

Théorèmes de dualité Quelque théorèmes établissent des dualités entre des espaces courants.

(1) Le théorème de représentation de Riesz 23.18 pour les espaces de Hilbert.

(2) La proposition 24.60 pour les espaces $L^p\big([0,1]\big)$ avec $1 < p < 2$.

(3) Le théorème de représentation de Riesz 24.64 pour les espaces L^p en général.

Tous ces théorèmes donnent la dualité par l'application $\Phi_x = \langle x, . \rangle$.

Thème 33 : opérations sur les distributions

(1) Convolution d'une distribution par une fonction, définition par l'équation (28.74).

(2) Dérivation d'une distribution, proposition-définition 28.15.

(3) Produit d'une distribution par une fonction, définition 28.14.

Thème 34 : transformée de Fourier

(1) Définition sur L^1, définition 27.2.

(2) La transformée de Fourier d'une fonction $L^1(\mathbb{R}^d)$ est continue, proposition 27.5.

(3) L'espace de Schwartz est stable par transformée de Fourier. L'application $\mathcal{F}\colon \mathscr{S}(\mathbb{R}^d) \to \mathscr{S}(\mathbb{R}^d)$ est une bijection linéaire et continue. Proposition 27.12

Thème 35 : convolution

(1) Définition pour $f, g \in L^1$, théorème 24.38.

(2) Inégalité de normes : si $f \in L^p$ et $g \in L^1$, alors $\|f * g\|_p \leqslant \|f\|_p \|g\|_1$, proposition 24.40.

(3) $\varphi \in L^1(\mathbb{R})$ et $\psi \in \mathscr{S}(\mathbb{R})$, alors $\varphi * \psi \in \mathscr{S}(\mathbb{R})$, proposition 24.87.

(4) Les suites régularisantes : $\lim_{n\to\infty} \rho_n * f = f$ dans la proposition 27.17.

(5) Convolution d'une distribution par une fonction, définition par l'équation (28.74).

Thème 36 : méthode de Newton

(1) Nous parlons un petit peu de méthode de Newton en dimension 1 dans 32.6.

(2) La méthode de Newton fonctionne bien avec les fonction convexes par la proposition 32.55.

(3) La méthode de Newton en dimension n est le théorème 32.61.

(4) Un intervalle de convergence autour de α s'obtient par majoration de $|g'|$, proposition 32.47.

(5) Un intervalle de convergence quadratique s'obtient par majoration de $|g''|$, théorème 32.53.

(6) En calcul numérique, section 32.6.

Thème 37 : méthodes de calcul

(1) Théorème de Rothstein-Trager 14.90.

(2) Algorithme des facteurs invariants 4.99.

(3) Méthode de Newton, théorème 32.61

(4) Calcul d'intégrale par suite équirépartie 26.8.

Thème 38 : définie positive

(1) Une application bilinéaire est définie positive lorsque $g(u,u) \geqslant 0$ et $g(u,u) = 0$ si et seulement si $u = 0$ est la définition 9.23.

(2) Un opérateur ou une matrice est défini positif si toutes ses valeurs propres sont positives, c'est la définition 9.167.

(3) Pour une matrice symétrique, définie positive implique $\langle Ax, x\rangle > 0$ pour tout x. C'est le lemme 9.170.

(4) Une application linéaire est définie positive si et seulement si sa matrice associée l'est. C'est la proposition 9.247.

Remarque : nous ne définissons pas la notion de matrice définie positive dans le cas d'une matrice non symétrique.

Thème 39 : norme matricielle et rayon spectral

(1) Définition du rayon spectral 10.10.

(2) Lien entre norme matricielle et rayon spectral, le théorème 10.25 assure que $\|A\|_2 = \sqrt{\rho(A^tA)}$.

(3) Lien entre valeurs propres et norme opérateur : le lemme 10.26 pour les matrices symétriques strictement définies positives donne $\|A\|_2 = \lambda_{max}$.

(4) Pour toute norme algébrique nous avons $\rho(A) \leqslant \|A\|$, proposition 10.15.

(5) Dans le cadre du conditionnement de matrice. Voir en particulier la proposition 32.104 qui utilise le théorème 10.25.

(6) Rayon spectral et convergence de méthode itérative, proposition 32.140.

Une norme matricielle donne une topologie. Il y a donc également des liens entre rayon spectral et convergence de série. Dans cette optique, pour les séries de matrices, voir le thème 40.

Thème 40 : série de matrices

(1) Rayon spectral et norme opérateur : thème 39.

(2) Exponentielle de matrices : thème 54.

(3) Série entière de matrices : section 15.12.

(4) Pour la série $\sum_k A^k = (1-A)^{-1}$.
— Pour un espace de Banach : proposition 10.52.
— Pour les matrices nilpotentes : proposition 9.138.
— En lien avec le rayon spectral (si et seulement si $\rho(A) < 1$) dans la proposition 15.90.
— Le lemme 15.44 parle de la série entière $\sum_{n\in\mathbb{N}} z^{nk} = (1-z^k)^{-1}$.
Cette série est utilisée entre autres dans la proposition 32.160 pour prouver qu'une M-matrice irréductible vérifie $A^{-1} > 0$.

Thème 41 : rang

(1) Définition 7.22.

(2) Le théorème du rang, théorème 7.23

(3) Prouver que des matrices sont équivalentes et les mettre sous des formes canoniques, lemme 7.28 et son corollaire 7.29.

(4) Tout hyperplan de $\mathbb{M}(n,\mathbb{K})$ coupe $\mathrm{GL}(n,\mathbb{K})$, corollaire 7.29. Cela utilise la forme canonique sus-mentionnée.

(5) Le lien entre application duale et orthogonal de la proposition 7.45 utilise la notion de rang.

(6) Prouver les équivalences à être un endomorphisme cyclique du théorème 9.233 via le lemme 9.232.

Thème 42 : extension de corps et polynômes

(1) Définition d'une extension de corps 5.47.

(2) Pour l'extension du corps de base d'un espace vectoriel et les propriétés d'extension des applications linéaires, voir la section 9.13.

(3) Extension de corps de base et similitude d'application linéaire (ou de matrices, c'est la même chose), théorème 9.235.

(4) Extension de corps de base et cyclicité des applications linéaires, corollaire 9.234.

(5) À propos d'extensions de $\mathbb{Q}$, le lemme 5.111.

Thème 43 : décomposition de matrices

(1) Décomposition de Bruhat, théorème 12.352.

(2) Décomposition de Dunford, théorème 9.194.

(3) Décomposition polaire 12.346 et la proposition 18.56 pour la régularité.

Thème 44 : systèmes d'équations linéaires

— Algorithme des facteurs invariants 4.99.
— La méthode du gradient à pas optimal permet de résoudre par itérations $Ax = b$ lorsque A est symétrique strictement définie positive. Il s'agit de minimiser une fonction bien choisit. Propositions 18.97 pour l'existence et 18.98 pour la méthode.

Thème 45 : racines de polynôme et factorisation de polynômes

(1) Si $\mathbb{A}$ est une anneau, la proposition 4.119 factorise une racine.

(2) Si $\mathbb{A}$ est un anneau, la proposition 4.121 factorise une racine avec sa multiplicité.

(3) Si $\mathbb{A}$ est un anneau, le théorème 4.123 factorise plusieurs racines avec leurs multiplicités.

(4) Si $\mathbb{K}$ est un corps et α une racine dans une extension, le polynôme minimal de α divise tout polynôme annulateur par la proposition 5.65.

(5) Le théorème 5.68 annule un polynôme de degré n ayant $n+1$ racines distinctes.

(6) La proposition 5.114 nous annule un polynôme à plusieurs variables lorsqu'il a trop de racines.

(7) En analyse complexe, le principe des zéros isolés 25.20 annule en gros toute série entière possédant un zéro non isolé.

(8) Polynômes irréductibles sur $\mathbb{F}_q$.

Thème 46 : formes bilinéaires et quadratiques

(1) Les formes bilinéaires ont été définies en 9.22.

(2) Forme quadratique, définition 9.242

Thème 47 : arithmétique modulo, théorème de Bézout

(1) Pour $\mathbb{Z}^*$ c'est le théorème 3.32.

(2) Théorème de Bézout dans un anneau principal : corollaire 4.77.

(3) Théorème de Bézout dans un anneau de polynômes : théorème 5.40.

(4) En parlant des racines de l'unité et des générateurs du groupe unitaire dans le lemme 3.99. Au passage nous y parlerons de solfège.

(5) La proposition 3.34 qui donne tout entier assez grand comme combinaisons de a et b à coefficients positifs est utilisée en chaînes de Markov, voir la définition 36.41 et ce qui suit.

(6) Calcul effectif du PGCD puis des coefficients de Bézout : sous-sections 3.5.4.1 et 3.5.4.2.

Thème 48 : polynômes

Coefficients dans un anneau commutatif (1) Les polynômes à coefficients dans un anneau commutatif sont à la section 4.10.

Coefficients dans un corps (1) Les polynômes à coefficients dans un corps sont à la section 5.3.

(2) Nous parlons de l'idéal des polynômes annulateurs dans le théorème 5.37.

(3) Le théorème 5.37 dit que $\mathbb{K}[X]$ est une anneau principal et que tous ses idéaux sont engendrés par un unique polynôme unitaire.

(4) Le polynôme minimal est irréductible, proposition 5.50.

Polynôme primitif (1) Un polynôme est irréductible sur A si et seulement si irréductible et primitif sur le corps des fractions, corollaire 4.124.

Thème 49 : invariants de similitude

(1) Théorème 9.221.

(2) Pour prouver que la similitude d'applications linéaires résiste à l'extension du corps de base, théorème 9.235.

(3) Pour prouver que la dimension du commutant d'un endomorphisme de E est de dimension au moins $\dim(E)$, lemme 9.232.

(4) Nous verrons dans la remarque 9.222 à propos des invariants de similitude que toute matrice est semblable à la matrice bloc-diagonale constituées des matrices compagnon (définition 9.216) de la suite des polynômes minimals.

Thème 50 : diagonalisation Des résultats qui parlent diagonalisation

(1) Définition d'un endomorphisme diagonalisable : 9.140.

(2) Conditions équivalentes au fait d'être diagonalisable en termes de polynôme minimal, y compris la décomposition en espaces propres : théorème 9.143.

(3) Diagonalisation simultanée 9.146, pseudo-diagonalisation simultanée 9.173.

(4) Diagonalisation d'exponentielle 9.197 utilisant Dunford.

(5) Décomposition polaire théorème 12.346. $M = SQ$, S est symétrique, réelle, définie positive, Q est orthogonale.

(6) Décomposition de Dunford 9.194. $u = s + n$ où s est diagonalisable et n est nilpotent, $[s, n] = 0$.

(7) Réduction de Jordan (bloc-diagonale) 9.224.

(8) L'algorithme des facteurs invariants 4.99 donne $U = PDQ$ avec P et Q inversibles, D diagonale, sans hypothèse sur U. De plus les éléments de D forment une chaîne d'éléments qui se divisent l'un l'autre.

Le théorème spectral et ses variantes :

(1) Théorème spectral, matrice symétrique, théorème 9.165. Via le lemme de Schur.

(2) Théorème spectral autoadjoint (c'est le même, mais vu sans matrices), théorème 9.261

(3) Théorème spectral hermitien, lemme 9.159.

(4) Théorème spectral, matrice normales, théorème 9.162.

Pour les résultats de décomposition dont une partie est diagonale, voir le thème 43 sur les décompositions.

Thème 51 : endomorphismes cycliques

(1) Définition 9.113.

(2) Son lien avec le commutant donné dans la proposition 9.230 et le théorème 9.233.

(3) Utilisation dans le théorème de Frobenius (invariants de similitude), théorème 9.221.

Thème 52 : déterminant

(1) Les n-formes alternées forment un espace de dimension 1, proposition 9.3.

(2) Déterminant d'une famille de vecteurs 9.6.

(3) Calcul d'un déterminant de taille 2×2 : équation (9.39).

(4) Déterminant d'un endomorphisme 9.9.

(5) Interprétations géométriques

   (a) À propos d'orthogonalité, le déterminant est très lié au produit vectoriel en dimension 3. Et il le généralise en dimension supérieure.

      i. Liaison au produit vectoriel (orthogonalité) dans la proposition 9.58.

      ii. En particulier le lemme 9.59 nous dit comment un déterminant donne un vecteur orthogonal à une famille donnée de vecteurs.

   (b) Déterminant et aires, volumes

      i. Déterminant et mesure de Lebesgue : théorème 13.223.

      ii. Aire du parallélogramme : proposition 16.49.

      iii. Volume du parallélépipède avec le produit mixte et le déterminant 3×3, 16.50.

Tant que nous en sommes dans les interprétations géométrique, il faut lier déterminant, produit vectoriel, orthogonalité et mesure en notant que l'élément de volume lors de l'intégration en dimension 3 est donné par (14.117) : $dS = \|T_u \times T_v\|$ qui est la norme du produit vectoriel des vecteurs tangents à la paramétrisation.

Nous voyons dans l'équation (14.114) que l'élément de volume pour une partie de dimension n dans $\mathbb{R}^m$ (à l'occasion d'y intégrer une fonction) est donné par un déterminant mettant en jeu les vecteurs tangents de la paramétrisation.

(6) Le déterminant de Vandermonde est à la proposition 9.12. Il est utilisé à divers endroits :
   (a) Pour prouver que $\mathrm{Tr}(u^p) = 0$ pour tout p si et seulement si u est nilpotente (lemme 9.136).
   (b) Pour prouver qu'un endomorphisme possédant $\dim(E)$ valeurs propres distinctes est cyclique (proposition 9.230).

Thème 53 : polynôme d'endomorphismes

(1) Endomorphismes cycliques et commutant dans le cas diagonalisable, proposition 9.230.
(2) Racine carré d'une matrice hermitienne positive, proposition 12.341.
(3) Théorème de Burnside sur les sous groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.238.
(4) Décomposition de Dunford, théorème 9.194.
(5) Algorithme des facteurs invariants 4.99.

Thème 54 : exponentielle de matrice

(1) Le lemme à propos d'exponentielle de matrice 15.94 donne :

$$\|e^{tA}\| \leqslant P\big(|t|\big) \sum_{i=1}^{r} e^{t\,\mathrm{Re}(\lambda_i)}. \tag{0.6}$$

(2) La proposition 9.197 : si $A \in \mathbb{M}(n, \mathbb{R})$ a un polynôme caractéristique scindé, alors A est diagonalisable si et seulement si e^A est diagonalisable.
(3) La section 15.12.3 parle des fonctions exponentielle et logarithme pour les matrices. Entre autres la dérivation et les séries.
(4) Pour résoudre des équations différentielles linéaires : sous-section 30.6.1.
(5) La proposition 9.196 dit que l'exponentielle est surjective sur $\mathrm{GL}(n, \mathbb{C})$.
(6) La proposition 10.89 : si u est un endomorphisme, alors $\exp(u)$ est un polynôme en u.
(7) Calcul effectif : sous-section 9.8.2.
(8) Proposition 12.338 : si $A \in \mathbb{M}(n, \mathbb{C})$ alors $e^{\mathrm{Tr}(A)} = \det(e^A)$.

Thème 55 : sous-groupes

(1) Théorème de Burnside sur les sous groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.238.
(2) Sous-groupes compacts de $\mathrm{GL}(n, \mathbb{R})$, lemme 12.353 ou proposition 12.354.

Thème 56 : groupe symétrique

(1) Définition 3.90.
(2) La table des caractères du groupe symétrique S_4 est donné dans la section 17.5.
(3) Le groupe symétrique S_4 est le groupe des symétries affines du tétraèdre régulier, proposition 16.17.
(4) Le groupe alterné A_5 est l'unique groupe simple d'ordre 60, proposition 8.49.
(5) La proposition 8.41 donne la position du groupe alterné dans le groupe symétrique : A_n est un sous-groupe caractéristique de S_n et l'unique sous-groupe d'indice 2.

Thème 57 : action de groupe

(1) Définition d'une action de groupe sur un ensemble : 3.72.

(2) Action du groupe modulaire sur le demi-plan de Poincaré, théorème 21.93.

(3) La formule de Burnside (théorème 3.85) parle du nombre d'orbites pour l'action d'un groupe fini sur un ensemble fini.

(4) Des applications de la formule de Burnside : le jeu de la roulette et l'affaire du collier, 16.4.4.1 et 16.4.4.2.

Thème 58 : classification de groupes

(1) Structure des groupes d'ordre pq, théorème 8.26.

(2) Le groupe alterné est simple, théorème 8.46.

(3) Définition 8.6 d'un p-groupe.

(4) Théorème de Sylow 8.11.

(5) Théorème de Burnside sur les sous groupes d'exposant fini de $\mathrm{GL}(n, \mathbb{C})$, théorème 9.238.

(6) $(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, corollaire 5.131.

Thème 59 : produit semi-direct de groupes

(1) Définition 3.93.

(2) Le corollaire 3.95 donne un critère pour prouver qu'un produit NH est un produit semi-direct.

(3) L'exemple 16.102 donne le groupe des isométries du carré comme un produit semi-direct.

(4) Le théorème 8.26 classifie les groupes d'ordre pq (p, q premiers distincts) à grands coups de produit semi-directs.

(5) Le théorème 16.6 donne les isométries de $\mathbb{R}^n$ par $\mathrm{Isom}(\mathbb{R}^n) = T(n) \times_\rho O(n)$ où $T(n)$ est le groupe des translations.

(6) La proposition 16.8 donne une décomposition du groupe orthogonal $\mathrm{O}(n) = \mathrm{SO}(n) \times_\rho C_2$ où $C_2 = \{\mathrm{Id}, R\}$ où R est de déterminant -1.

(7) La proposition 11.53 donne $\mathrm{Aff}(\mathbb{R}^n) = T(n) \times_\rho \mathrm{GL}(n, \mathbb{R})$ où $\mathrm{Aff}(\mathbb{R}^n)$ est le groupe des applications affines bijectives de $\mathbb{R}^n$.

Thème 60 : théorie des représentations

(1) Table des caractères du groupe diédral, section 17.6.

(2) Table des caractères du groupe symétrique S_4, section 17.5.

Thème 61 : isométries Ne pas confondre une isométrie d'un espace affine avec une isométrie d'un espace euclidien. Les isométrie d'un espace euclidien préservent le produit scalaire et fixent donc l'origine (lemme 9.28). Les isométrie des espaces affines par contre conservent les distances (définition 11.54) et peuvent donc déplacer l'origine de l'espace vectoriel sur lequel il est modelé ; typiquement les translations sont des isométries de l'espace affine mais pas de l'espace euclidien.

Parfois, lorsqu'on coupe les cheveux en quatre, il faut faire attention en parlant de $\mathbb{R}^n$: soit on en parle comme d'un espace métrique (muni de la distance), soit on en parle comme d'un espace normé (muni de la norme ou du produit scalaire).

(1) Définition d'une isométrie pour une forme bilinéaire, 9.27. Pour une forme quadratique : définition 16.2.

(2) Définition du groupe orthogonal 9.48, et le spécial orthogonal $\mathrm{SO}(n)$ en la définition 9.51. Le groupe $\mathrm{SO}(2)$ est le groupe des rotations, par corollaire 16.81.

(3) Le lemme 16.113 donne à toute rotation une matrice de la forme connue. C'est autour de cela que nous définissons les angles.

(4) Le groupe orthogonal est le groupe des isométries de $\mathbb{R}^n$, proposition 9.50.

(5) Les isométries de l'espace euclidien sont affines, 16.4.

(6) Les isométries de l'espace euclidien comme produit semi-direct : $\mathrm{Isom}(\mathbb{R}^n) \simeq T(n) \times_\rho \mathrm{O}(n)$, théorème 16.6.

(7) Isométries du cube, section 10.3.

(8) Générateurs du groupe diédral, proposition 16.100.

Thème 62 : caractérisation de distributions en probabilités

(1) La probabilité conjointe est la définition 34.19.

(2) La fonction de répartition est la définition 34.53.

(3) La fonction caractéristique est la définition 34.55.

Thème 63 : théorème central limite

(1) Pour les processus de Poisson, théorème 38.5.

Thème 64 : lemme de transfert Il s'agit du résultat $\hat{f}' = i\xi\hat{f}$.

(1) Lemme 27.11 sur $\mathscr{S}(\mathbb{R}^d)$

(2) Lemme 29.12 pour L^2.

Thème 65 : probabilités et espérances conditionnelles Les deux définitions de base, sur lesquelles se basent toutes les choses conditionnelles sont :

— La probabilité conditionnelle d'un événement en sachant un autre : $P(A|B)$ de la définition 34.29.

— L'espérance conditionnelle d'une variable aléatoire sachant une tribu : $E(X|\mathcal{F})$ de la définition 34.31.

Les autres sont listées ci-dessous.

La probabilité conditionnelle d'un événement par rapport à un autre donnée dans la proposition 34.29 est le nombre

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{0.7}$$

La probabilité conditionnelle d'un événement vis-à-vis d'une variable aléatoire discrète est par la définition 34.47 la variable aléatoire donnée par

$$P(A|X)(\omega) = P(A|X = X(\omega)). \tag{0.8}$$

Dans le cas continu, c'est la définition 34.48 :

$$P(A|X) = P(A|\sigma(X)) = E(\mathbb{1}_A|\sigma(X)). \tag{0.9}$$

L'espérance conditionnelle d'une variable aléatoire par rapport à une tribu $E(X|\mathcal{F})$ est la variable aléatoire $\mathcal{F}$-mesurable telle que

$$\int_B E(X|\mathcal{F}) = \int_B X \tag{0.10}$$

pour tout $X \in \mathcal{F}$. Si $X \in L^2(\Omega, \mathcal{A}, P)$ alors $E(X|\mathcal{F}) = \mathtt{proj}_K(X)$ où K est le sous-ensemble de $L^2(\Omega, \mathcal{A}, P)$ des fonctions $\mathcal{F}$-mesurables (théorème 34.31). Cela au sens des projections orthogonales.

La probabilité conditionnelle d'un événement par rapport à une tribu est la variable aléatoire

$$P(A|\mathcal{F}) = E(\mathbb{1}_A|\mathcal{F}). \tag{0.11}$$

L'espérance conditionnelle d'une variable aléatoire par rapport à une autre de la définition 34.33 est une variation sur le thème :

$$E(X|Y) = E(X|\sigma(Y)), \tag{0.12}$$

Notons que partout, si X est une variable aléatoire, la notation «sachant X» est un raccourcis pour dire «sachant la tribu engendrée par X».

Thème 66 : dénombrements

— Coloriage de roulette (16.4.4.1) et composition de colliers (16.4.4.2).
— Nombres de Bell, théorème 15.69.
— Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décomposition de fractions en éléments simples, théorème 15.47.

Thème 67 : enveloppes

(1) L'ellipse de John-Loewner donne un ellipsoïde de volume minimum autour d'un compact dans $\mathbb{R}^n$, théorème 18.100.
(2) Le cercle circonscrit à une courbe donne un cercle de rayon minimal contenant une courbe fermée simple, proposition 19.83.
(3) Enveloppe convexe du groupe orthogonal 12.351.
(4) Enveloppe convexe d'une courbe fermée plane comme intersection des demi-plans tangents, proposition 19.89.

Thème 68 : équations diophantiennes

(1) Équation $ax + by = c$ dans $\mathbb{N}$, équation (3.74).
(2) Dans 3.5.7, nous résolvons $ax + by = c$ en utilisant Bézout (théorème 3.32).
(3) L'exemple 4.89 donne une application de la pure notion de modulo pour $x^2 = 3y^2 + 8$. Pas de solutions.
(4) L'exemple 4.90 résout l'équation $x^2 + 2 = y^3$ en parlant de l'extension $\mathbb{Z}[i\sqrt{2}]$ et de stathme.
(5) Les propositions 4.94 et 4.96 parlent de triplets pythagoriciens.
(6) Le dénombrement des solutions de l'équation $\alpha_1 n_1 + \dots \alpha_p n_p = n$ utilise des séries entières et des décomposition de fractions en éléments simples, théorème 15.47.
(7) La proposition 2.18 donne une bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ en résolvant dans $\mathbb{N}$ (entre autres) l'équation $k = y^2 + x$ pour k fixé.

Table des matières

Thématique **2**

Table des matières **19**

Index **46**

Liste des notations **68**

1 Introduction (Vol 1) **73**
1.1 Préface . . . 73
1.2 Auteurs, contributeurs, sources et remerciements . . . 73
1.2.1 Ceux qui ont travaillé sur le Frido . . . 73
1.2.2 Aide directe, mais pas volontairement sur le Frido . . . 74
1.2.3 Des gens qui ont fait un travail qui m'a bien servi . . . 74
1.3 Les choses qui doivent vous faire tiquer . . . 75
1.4 Les questions pour lesquelles je n'ai pas (encore) de réponse . . . 75
1.4.1 Mes questions d'algèbre, géométrie. . . . 75
1.4.2 Mes questions de probabilité et statistiques. . . . 77
1.4.3 Les preuves à relire par des experts . . . 77
1.5 Comment contribuer et aider (math) ? . . . 82
1.5.1 Niveau facile . . . 82
1.5.1.1 Géométrie . . . 83
1.5.2 Niveau moyen . . . 83
1.5.2.1 Questions d'analyse . . . 83
1.5.2.2 Questions de géométrie . . . 84
1.5.2.3 Questions de numérique . . . 84
1.5.3 Niveau difficile . . . 84
1.5.3.1 Questions d'analyse . . . 84
1.5.4 Question de numérique . . . 86
1.5.5 Pour ma culture générale . . . 86
1.6 Comment contribuer et aider (LaTeX et programmation) . . . 86
1.6.1 Niveau facile . . . 86
1.6.2 Niveau moyen . . . 86
1.6.3 Niveau difficile . . . 86
1.7 Sage est là pour vous aider . . . 87
1.7.1 Lancez-vous dans Sage . . . 87
1.7.2 Exemples de ce que Sage peut faire pour vous . . . 87

2 Construction des ensembles de nombres (Vol 1) **89**
2.1 Quelques éléments sur les ensembles . . . 89
2.1.1 Ensemble ordonné . . . 89
2.1.2 Lemme de Zorn . . . 90
2.1.3 Complémentaire . . . 90
2.1.4 Relations d'équivalence . . . 91

2.2 Les naturels . . . 92
2.2.1 La construction . . . 92
2.2.2 Quelque autres résultats . . . 92
2.3 Les entiers . . . 93
2.4 Quelques structures algébriques . . . 93
2.5 Les rationnels . . . 95
2.5.1 Suites de Cauchy dans les rationnels . . . 95
2.5.2 Insuffisance des rationnels . . . 97
2.6 Les réels . . . 100
2.6.1 L'ensemble . . . 100
2.6.2 Relation d'ordre . . . 102
2.7 Les complexes . . . 105
2.7.1 Définitions . . . 106

3 Théorie des groupes (Vol 1) 107
3.1 Groupes . . . 107
3.2 Sous groupe normal . . . 108
3.2.1 Classes de conjugaison . . . 109
3.3 Groupe dérivé . . . 109
3.4 Théorèmes d'isomorphismes . . . 110
3.5 Le groupe et anneau des entiers . . . 112
3.5.1 Division euclidienne . . . 112
3.5.2 Sous-groupes de $(\mathbb{Z}, +)$ . . . 113
3.5.3 PGCD, PPCM et Bézout . . . 113
3.5.4 Calcul effectif du PGCD et de Bézout . . . 116
3.5.4.1 Algorithme d'Euclide pour le PGCD . . . 116
3.5.4.2 Algorithme étendu : calcul effectif des coefficients de Bézout . . . 117
3.5.5 Décomposition en facteurs premiers . . . 118
3.5.6 Écriture des fractions . . . 119
3.5.7 Équation diophantienne linéaire à deux inconnues . . . 120
3.5.8 Quotients . . . 121
3.6 Indice d'un sous-groupe et ordre des éléments . . . 122
3.7 Suite de composition . . . 125
3.8 Groupes résolubles . . . 128
3.9 Action de groupes . . . 129
3.10 Groupe des permutations . . . 134
3.11 Produit semi-direct de groupes . . . 134
3.12 Groupe de torsion . . . 136
3.13 Famille presque nulle . . . 136
3.14 Le groupe des racines de l'unité . . . 136
3.15 Fonction indicatrice d'Euler (première partie) . . . 138
3.15.1 Introduction par les racines de l'unité . . . 139
3.15.2 Générateurs . . . 140

4 Anneaux (Vol 1) 141
4.1 Généralités . . . 141
4.2 Binôme de Newton et morphisme de Frobenius . . . 143
4.3 Idéal dans un anneau . . . 144
4.3.1 Résultats supplémentaires sur l'anneau des entiers . . . 145
4.4 Caractéristique . . . 145
4.5 Modules . . . 146
4.6 Anneau intègre . . . 148
4.6.1 Caractéristique d'un anneau intègre . . . 149
4.6.2 Divisibilité et classes d'association . . . 150

4.6.3 PGCD et PPCM . . . 150
4.6.4 Anneaux intègres et corps . . . 151
4.6.5 Corps des fractions . . . 152
4.7 Anneau factoriel . . . 152
4.8 Anneau principal . . . 153
4.8.1 Bézout . . . 154
4.8.2 Anneau noetherien . . . 156
4.9 Anneau euclidien . . . 156
4.9.1 Équations diophantiennes . . . 158
4.9.2 Triplets pythagoriciens et équation de Fermat pour $n = 4$ . . . 159
4.9.3 Lignes et colonnes de matrices . . . 162
4.9.4 Algorithme des facteurs invariants . . . 163
4.10 Polynômes à coefficients dans un anneau commutatif . . . 165
4.10.1 Division euclidienne . . . 166
4.10.2 Polynôme primitif . . . 167
4.10.3 Racines des polynômes . . . 167
4.10.4 Quelques identités . . . 169

5 Corps (Vol 1) 171
5.1 Généralités . . . 171
5.1.1 Corps ordonnés . . . 171
5.1.2 Automorphismes de $\mathbb{R}$ et $\mathbb{C}$ . . . 171
5.1.3 Corps premier . . . 173
5.1.4 Petit théorème de Fermat . . . 173
5.2 Théorème des deux carrés . . . 174
5.2.1 Un peu de structure dans $\mathbb{Z}[i]$ . . . 174
5.2.2 Résultats chinois . . . 178
5.3 Polynômes à coefficients dans un corps . . . 180
5.3.1 Irréductibilité . . . 180
5.3.2 Idéaux . . . 181
5.3.3 Bézout . . . 183
5.3.4 Lemme et théorème de Gauss . . . 183
5.4 Extension de corps . . . 184
5.4.1 Polynôme minimal . . . 184
5.4.2 Racines de polynômes . . . 187
5.4.3 Corps de rupture . . . 189
5.4.4 Corps de décomposition . . . 191
5.4.5 Clôture algébrique . . . 192
5.4.6 Extensions séparables . . . 192
5.4.7 Idéal maximum . . . 196
5.5 Espaces de polynômes . . . 198
5.5.1 Polynômes symétriques, alternés ou semi-symétriques . . . 200
5.5.2 Polynôme symétrique élémentaire . . . 200
5.5.3 Relations coefficients racines . . . 202
5.6 Polynômes cyclotomiques . . . 203
5.6.1 Définitions et propriétés . . . 203
5.6.2 Nombres premiers . . . 207
5.6.3 Théorème de Wedderburn . . . 209
5.7 Corps finis . . . 211
5.7.1 Existence, unicité . . . 211
5.7.2 Symboles de Legendre et carrés . . . 213
5.7.3 Théorème de Chevalley-Warning . . . 219
5.7.4 Contenu d'un polynôme . . . 221
5.7.5 Théorème de l'élément primitif . . . 222

5.7.6 Construction de $\mathbb{F}_q$ . . . 226
5.7.6.1 La version du faignant . . . 226
5.7.6.2 La version plus élaborée . . . 227
5.7.7 Exemple : étude de $\mathbb{F}_{16}$ . . . 229
5.7.8 Polynômes irréductibles sur $\mathbb{F}_q$ . . . 231
5.8 Constructions à la règle et au compas . . . 233
5.8.1 Quelque constructions . . . 233
5.8.2 Nombres constructibles . . . 235
5.8.3 Polygones constructibles . . . 237
5.8.4 Matrices . . . 242
5.9 Minuscule morceau sur la théorie de Galois . . . 242
5.9.1 Théorème de d'Alembert . . . 243

6 Topologie générale (Vol 1) 245
6.1 Topologie en général . . . 245
6.1.1 Définitions basiques . . . 245
6.1.2 Topologie produit . . . 246
6.1.3 Séparabilité . . . 246
6.1.4 Topologie métrique . . . 247
6.1.5 Espace vectoriel topologique . . . 249
6.2 Espace vectoriel normé . . . 249
6.2.1 Introduction : norme et valeur absolue . . . 249
6.2.2 Norme . . . 250
6.2.3 Quelque exemples . . . 251
6.3 Suites de Cauchy, métrique et espaces complet . . . 252
6.3.1 Généralités . . . 252
6.3.2 Cas d'équivalence . . . 254
6.3.3 Espace topologique normé . . . 256
6.3.4 Espace métrique . . . 256
6.4 Topologie sur l'ensemble des réels . . . 257
6.5 Base de topologie . . . 259
6.6 Limite et continuité . . . 260
6.7 Compacité . . . 261
6.7.1 Quelque propriétés . . . 261
6.8 Topologie induite . . . 263
6.9 Connexité . . . 264
6.10 Un peu de topologie réelle . . . 266
6.10.1 Suites numériques . . . 266
6.10.2 Maximum, supremum et compagnie . . . 268
6.10.2.1 ... et quelque exemples . . . 268
6.10.3 Intervalles . . . 273
6.10.3.1 Connexité . . . 273
6.10.3.2 Compacité . . . 274
6.10.4 Image d'un compact . . . 276
6.10.5 Connexité par arcs . . . 278
6.10.6 Topologie de la droite réelle complétée . . . 278
6.10.7 Limite pointée ou épointée ? . . . 279
6.11 Espaces métriques . . . 279
6.11.1 Espaces métrisables . . . 279
6.11.2 Fonctions continues . . . 279
6.11.3 Caractérisations séquentielles . . . 280
6.11.4 Distance entre un point et un ensemble . . . 282
6.11.5 Compacité . . . 283
6.11.6 Ensembles enchaînés . . . 288

6.11.7 Produit fini d'espaces métriques . . . 289
6.12 Ensembles nulle part denses . . . 290
6.13 Encore de la topologie réelle . . . 291
6.13.1 Ouverts et fermés . . . 291
6.13.2 Intérieur, adhérence et frontière . . . 291
6.13.3 Point d'accumulation, point isolé . . . 293
6.13.4 Limite de suite . . . 293
6.14 Application réciproque . . . 294
6.15 Topologie des semi-normes . . . 295
6.15.1 Espace dual . . . 297
6.15.2 Espace $C^k(\mathbb{R}, E')$ . . . 297
6.16 Espaces de Baire . . . 298

7 Espaces vectoriels (Vol 1) 301
7.1 Parties libres, génératrices, bases et dimension . . . 301
7.2 Applications linéaires . . . 305
7.2.1 Définition . . . 305
7.2.2 Rang . . . 306
7.2.3 Matrices équivalentes et similaires . . . 307
7.3 Matrice d'une application linéaire . . . 309
7.3.1 Écriture dans une base . . . 309
7.3.2 Changement de base : vecteurs de base . . . 310
7.3.3 Changement de base : coordonnées . . . 311
7.3.4 Changement de base : matrice d'une application linéaire . . . 311
7.3.5 Changement de base : matrice d'une forme bilinéaire . . . 312
7.4 Dualité . . . 312
7.4.1 Orthogonal . . . 312
7.4.2 Transposée . . . 313
7.4.3 Transposée : sans le dual . . . 315
7.4.4 Polynômes de Lagrange . . . 316
7.4.5 Dual de $\mathbb{M}(n, \mathbb{K})$ . . . 316

8 Classification de certains groupes (Vol 1) 319
8.1 Théorèmes de Sylow . . . 319
8.2 Un peu de classification de groupes . . . 323
8.2.1 Groupe monogène . . . 323
8.2.2 Automorphismes du groupe $\mathbb{Z}/n\mathbb{Z}$ . . . 324
8.2.3 Groupes abéliens finis . . . 326
8.2.4 Groupes d'ordre pq . . . 328
8.2.5 Fonction indicatrice d'Euler . . . 331
8.3 Le groupe symétrique . . . 332
8.3.1 Décomposition en cycles . . . 332
8.3.2 Le groupe alterné . . . 334
8.3.3 Sous-groupes normaux . . . 338
8.3.4 Indice . . . 340

9 Suite des espaces vectoriels (Vol 1) 343
9.1 Déterminants . . . 343
9.1.1 Formes multilinéaires alternées . . . 343
9.1.2 Déterminant d'une famille de vecteurs . . . 345
9.1.3 Déterminant d'un endomorphisme . . . 347
9.1.4 Déterminant d'une matrice . . . 349
9.1.5 Déterminant de Vandermonde . . . 349
9.1.6 Déterminant de Gram . . . 352

9.1.7 Déterminant de Cauchy . . . 352
9.1.8 Matrice de Sylvester . . . 352
9.1.9 Théorème de Kronecker . . . 355
9.2 Produit scalaire . . . 357
9.2.1 Projection et angles . . . 359
9.2.2 Procédé de Gram-Schmidt . . . 361
9.3 Hermitien, orthogonal, adjoint . . . 362
9.3.1 Opérateur orthogonal, matrice orthogonale . . . 363
9.4 Directions conservées . . . 364
9.4.1 Généralités . . . 364
9.4.2 Comment trouver la matrice d'une symétrie donnée ? . . . 365
9.4.2.1 Symétrie par rapport à un plan . . . 365
9.4.2.2 Symétrie par rapport à une droite . . . 366
9.4.2.3 En résumé . . . 367
9.5 Produit vectoriel . . . 367
9.5.1 Produit vectoriel . . . 368
9.5.2 Produit mixte . . . 369
9.6 Topologie . . . 371
9.6.1 Boules et sphères . . . 371
9.6.2 Ouverts, fermés, intérieur et adhérence . . . 372
9.6.3 Point isolé, point d'accumulation . . . 379
9.7 Valeur propre et vecteur propre . . . 380
9.8 Polynôme d'endomorphismes . . . 380
9.8.1 Polynômes d'endomorphismes . . . 381
9.8.2 Calcul effectif de l'exponentielle d'une matrice . . . 383
9.8.3 Polynôme minimal et minimal ponctuel . . . 383
9.8.4 Polynôme caractéristique . . . 389
9.9 Diagonalisation et trigonalisation . . . 393
9.9.1 Matrices semblables . . . 393
9.9.2 Endomorphismes nilpotents . . . 394
9.9.3 Endomorphismes diagonalisables . . . 397
9.9.4 Diagonalisation : cas complexe, pas toujours . . . 400
9.9.5 Trigonalisation : généralités . . . 400
9.9.6 Trigonalisation : cas complexe . . . 401
9.9.7 Diagonalisation : cas complexe, ce qu'on a . . . 403
9.9.8 Diagonalisation : cas réel . . . 405
9.9.9 Pseudo-réduction simultanée . . . 408
9.10 Sommes de familles infinies . . . 409
9.10.1 Convergence commutative . . . 409
9.11 Fonctions . . . 413
9.12 Sous espaces caractéristiques . . . 414
9.12.1 Théorèmes de décomposition . . . 415
9.12.2 Diverses conséquences . . . 417
9.12.3 Diagonalisabilité d'exponentielle . . . 418
9.12.4 Valeurs singulières . . . 420
9.13 Extension du corps de base . . . 420
9.13.1 Extension des applications linéaires . . . 420
9.13.2 Projections . . . 422
9.13.3 Rang, polynôme minimal, polynôme caractéristique . . . 425
9.14 Frobenius et Jordan . . . 426
9.14.1 Matrice compagnon . . . 426
9.14.2 Réduction de Frobenius . . . 427
9.14.3 Forme normale de Jordan . . . 430

9.15 Commutant et endomorphismes cycliques . . . 431
9.15.1 Endomorphisme cyclique . . . 431
9.15.2 Commutant : cas diagonalisable . . . 432
9.15.3 Commutant : cas général . . . 434
9.16 Théorème de Burnside . . . 437
9.16.1 Théorème de Lie-Kolchin . . . 438
9.17 Formes bilinéaires et quadratiques . . . 441
9.17.1 Généralités . . . 441
9.17.2 Topologie . . . 441
9.17.3 Matrice associée . . . 441
9.17.4 Dégénérescence . . . 443
9.17.5 Inégalité de Minkowski . . . 444
9.17.6 Ellipsoïde . . . 445
9.18 Théorème spectral auto-adjoint . . . 446
9.19 Mini introduction au produit tensoriel . . . 449
9.19.1 Définitions . . . 449
9.19.2 Application d'opérateurs . . . 449
9.20 Méthode de Gauss pour résoudre des systèmes d'équations linéaires . . . 450

10 Espaces vectoriels normés (Vol 1) 451
10.1 Équivalence des normes . . . 451
10.1.1 En dimension finie . . . 451
10.1.2 Contre-exemple en dimension infinie . . . 452
10.2 Norme opérateur . . . 453
10.2.1 Norme d'algèbre . . . 455
10.2.2 Application linéaire continue et bornée . . . 459
10.2.3 Normes de matrices et d'applications linéaires . . . 461
10.3 Isométriques du cube . . . 462
10.4 Chiffrement RSA . . . 464
10.4.1 Mise en place par Bob . . . 464
10.4.2 Chiffrement . . . 464
10.4.3 Déchiffrement . . . 465
10.4.4 Une imprudence à ne pas commettre . . . 465
10.4.5 Problèmes calculatoires . . . 466
10.4.6 La solidité de RSA . . . 466
10.4.7 Note non mathématique pour doucher l'enthousiasme . . . 466
10.5 Suites . . . 466
10.5.1 Limites, convergence . . . 466
10.5.2 Critère de Cauchy . . . 467
10.5.3 Approximation . . . 468
10.6 Séries . . . 469
10.7 Série réelle . . . 472
10.7.1 Critères de convergence absolue . . . 472
10.7.2 Critères de convergence simple . . . 474
10.7.2.1 Critère d'Abel . . . 474
10.7.3 Quelques séries usuelles . . . 475
10.7.4 Moyenne de Cesaro . . . 476
10.7.5 Écriture décimale d'un nombre . . . 476
10.8 Produit fini d'espaces vectoriels normés . . . 480
10.8.1 Norme . . . 480
10.8.2 Suites . . . 482
10.8.3 Continuité du produit de matrices . . . 483
10.9 Applications multilinéaires . . . 484
10.10Exponentielle de matrice . . . 486

10.11 Mini introduction aux nombres p-adiques . . . 487
10.11.1 La flèche d'Achille . . . 487
10.11.2 La tortue et Achille . . . 487
10.11.3 Dans les nombres p-adiques, c'est vrai . . . 488

11 Espaces affines (Vol 2) 491
11.1 Repères affines . . . 491
11.2 Classification affine des conique . . . 492
11.3 Applications affines . . . 494
11.4 Isomorphismes . . . 495
11.5 Sous espaces affines . . . 496
11.6 Barycentre . . . 497
11.6.1 Sous-espaces affines . . . 498
11.6.2 Enveloppe convexe . . . 500
11.6.3 Applications affines et barycentre . . . 503
11.7 Repères, coordonnées cartésiennes et barycentriques . . . 504
11.7.1 Équation de droite . . . 506
11.7.2 Associativité, coordonnées barycentriques dans un triangle . . . 506
11.8 Structure du groupe des applications affines . . . 507
11.9 Isométries . . . 509

12 Analyse réelle (Vol 2) 511
12.1 Intervalles . . . 511
12.2 Application réciproque . . . 512
12.2.1 Définitions . . . 512
12.2.2 Graphe de la fonction réciproque . . . 513
12.3 Fonctions continues . . . 514
12.3.1 La fonction la moins continue du monde . . . 516
12.3.2 Approche topologique . . . 516
12.3.3 Continuité de la racine carré, invitation à la topologie induite . . . 518
12.3.4 Limites en des nombres . . . 520
12.3.5 Limites quand tout va bien . . . 520
12.3.6 Discussion avec mon ordinateur . . . 525
12.3.7 Limites et prolongement . . . 525
12.3.8 Théorèmes des valeurs intermédiaires . . . 526
12.3.9 Théorème de la bijection . . . 526
12.3.10 Prolongement par continuité . . . 529
12.4 Espace des fonctions continues . . . 530
12.5 Uniforme continuité . . . 535
12.6 Compacité . . . 537
12.7 Limite et continuité . . . 539
12.7.1 Définition . . . 539
12.7.2 Propriétés de base . . . 540
12.7.3 Limites de fonctions . . . 540
12.7.4 Règles simples de calcul . . . 546
12.8 Limites à plusieurs variables . . . 546
12.8.1 Règle de l'étau . . . 547
12.8.2 Méthode des chemins . . . 548
12.8.3 Méthode des coordonnées polaires . . . 550
12.8.4 Méthode du développement asymptotique . . . 554
12.9 Dérivée : exemples introductifs . . . 554
12.9.1 La vitesse . . . 554
12.9.2 La tangente à une courbe . . . 555
12.9.3 L'aire en dessous d'une courbe . . . 556

12.10Continuité et dérivabilité . . . 557
12.11Dérivation de fonctions d'une variable réelle . . . 558
12.11.1 Exemples . . . 558
12.11.1.1 La fonction $f(x) = x$ . . . 558
12.11.1.2 La fonction $f(x) = x^2$ . . . 558
12.11.1.3 La fonction $f(x) = \sqrt{x}$ . . . 559
12.11.2 Calcul de la dérivée . . . 559
12.11.3 Interprétation géométrique : tangente . . . 560
12.11.4 Interprétation géométrique : approximation affine . . . 560
12.12Opérations sur les dérivées . . . 561
12.12.1 Développement limité au premier ordre . . . 562
12.13Dérivation à une variable . . . 563
12.13.1 Définition . . . 563
12.13.2 Quelques formules à connaître . . . 563
12.14Dérivation et croissance . . . 563
12.14.1 Théorèmes de Rolle et des accroissements finis . . . 565
12.14.2 Règle de l'Hospital . . . 567
12.14.3 À propos de primitives . . . 569
12.15Fonctions de plusieurs variables . . . 570
12.15.1 Graphes de fonctions à plusieurs variables . . . 573
12.15.2 Courbes de niveau . . . 574
12.16Dérivée directionnelle . . . 578
12.16.1 Dérivée partielle et directionnelles . . . 578
12.16.1.1 Quelques propriétés et notations . . . 579
12.16.2 Gradient : direction de plus grande pente . . . 583
12.16.3 Gradient : orthogonal au plan tangent . . . 583
12.17Dérivée directionnelle de fonctions composées . . . 584
12.18Formes différentielles . . . 586
12.18.1 Décomposition dans la base duale . . . 586
12.18.2 L'isomorphisme musical . . . 587
12.18.3 Différentielles de fonctions composées . . . 587
12.18.4 Exemple de composée : les coordonnées polaires . . . 588
12.19Différentielle . . . 589
12.19.1 Exemples introductifs . . . 590
12.19.2 Définition de la différentielle . . . 591
12.19.3 Base duale . . . 592
12.19.4 Pas la différentielle extérieure . . . 592
12.19.5 Continuité, dérivabilité et différentiabilité . . . 593
12.19.6 Calcul de valeurs approchées . . . 595
12.19.7 Prouver qu'un fonction n'est pas différentiable . . . 598
12.19.7.1 Continuité . . . 598
12.19.7.2 Linéarité . . . 598
12.19.7.3 Cohérence des dérivées partielles et directionnelle . . . 601
12.19.7.4 Un candidat dans la définition (marche toujours) . . . 602
12.19.8 Règles de calcul . . . 602
12.19.9 Linéarité . . . 603
12.19.10 Produit . . . 603
12.19.11 Différentielle de fonction composée . . . 605
12.19.12 Différentielle et dérivées partielles . . . 608
12.19.13 Plan tangent . . . 609
12.19.14 Calcul de différentielles . . . 609
12.19.15 Notes idéologiques quant au concept de plan tangent . . . 610
12.19.16 Gradient et recherche du plan tangent . . . 610

12.20 Calcul différentiel dans un espace vectoriel normé . . . 612
12.20.1 Différentielle . . . 612
12.20.2 (non ?) Différentiabilité des applications linéaires . . . 613
12.20.3 Dérivation en chaine et formule de Leibnitz . . . 613
12.20.4 Différentielle partielle . . . 617
12.20.5 Formule des accroissements finis . . . 618
12.20.6 L'inverse, sa différentielle . . . 621
12.20.7 Projection orthogonale . . . 623
12.21 Jacobienne . . . 624
12.21.1 Rappels et définitions . . . 624
12.22 Fonctions de classe $\mathcal{C}^1$ . . . 625
12.23 Théorèmes des accroissements finis . . . 626
12.24 Fonctions Lipschitziennes . . . 627
12.25 Différentielles d'ordre supérieur . . . 629
12.25.1 Identification des espaces d'applications multilinéaires . . . 629
12.25.2 Fonctions différentiables plusieurs fois . . . 629
12.25.3 Différentielle seconde, fonction de classe C^2 . . . 630
12.25.4 Ordre supérieur . . . 633
12.26 Développement asymptotique, théorème de Taylor . . . 634
12.26.1 Fonctions «petit o» . . . 635
12.26.2 Autres formulations . . . 636
12.26.3 Formule et reste . . . 637
12.26.4 Reste intégral . . . 637
12.27 Développement limité autour de zéro . . . 637
12.27.1 Généralités . . . 638
12.27.2 Formule de Taylor-Young . . . 639
12.27.3 Règles de calcul . . . 641
12.27.3.1 Linéarité des développements limités . . . 641
12.27.3.2 Développement limité d'un quotient . . . 642
12.27.3.3 Développement limité d'une fonction composée . . . 643
12.28 Développement au voisinage de $x_0 \neq 0$ . . . 644
12.29 Application au calcul de limites . . . 645
12.30 Développement au voisinage de l'infini . . . 646
12.31 Convergence uniforme . . . 646
12.31.1 Critère de Cauchy uniforme . . . 646
12.31.2 Complétude avec la norme uniforme . . . 647
12.31.3 Permuter avec les dérivées partielles . . . 649
12.32 Fonctions réelles de deux variables réelles . . . 650
12.32.1 Limites de fonctions à deux variables . . . 650
12.32.2 Dérivées partielles . . . 651
12.32.3 Différentielle et accroissement . . . 652
12.33 Les fonctions à valeurs vectorielles . . . 652
12.34 Fonctions vectorielles de plusieurs variables . . . 653
12.35 Champs de vecteurs . . . 653
12.35.1 Matrice jacobienne . . . 654
12.36 Divergence, rotationnel et l'opérateur nabla . . . 654
12.37 Interprétation de la divergence . . . 657
12.38 Quelques formules de Leibnitz . . . 659
12.39 Action de groupe et connexité . . . 659
12.40 Espaces de matrices . . . 661
12.40.1 Dilatations et transvections . . . 662
12.40.2 Connexité de certains groupes . . . 667
12.40.3 Densité . . . 669

12.40.4 Racine carré d'une matrice hermitienne positive . . . 671
12.40.5 Racine carré d'une matrice symétrique positive . . . 672
12.40.6 Décomposition polaires : cas réel . . . 673
12.40.7 Enveloppe convexe . . . 675
12.40.8 Décomposition de Bruhat . . . 678
12.41 Sous-groupes du groupe linéaire . . . 680

13 Tribus, théorie de la mesure (Vol 2) 685
13.1 Tribus . . . 685
13.1.1 Généralités . . . 685
13.1.1.1 Tribu induite . . . 686
13.1.1.2 Tribu borélienne . . . 687
13.1.1.3 Les boréliens de $\mathbb{R}$ . . . 687
13.1.2 Tribu de Baire . . . 687
13.2 Théorie de la mesure . . . 689
13.2.1 Mesure et algèbre de parties . . . 690
13.2.2 Mesure sur un espace mesurable . . . 692
13.2.3 Mesure extérieure . . . 698
13.2.4 Espace mesuré complet . . . 700
13.2.5 Prolongement . . . 708
13.3 Applications mesurables . . . 710
13.3.1 Propriétés . . . 710
13.3.2 D'une tribu à l'autre . . . 710
13.3.3 Mesure image . . . 713
13.3.4 Régularité d'une mesure . . . 714
13.3.5 Théorème de récurrence . . . 718
13.4 Mesurabilité des fonctions à valeurs réelles . . . 719
13.4.1 Quelques mots à propos de $\overline{\mathbb{R}}$ . . . 719
13.4.2 Limite supérieure et inférieure . . . 719
13.4.3 Fonctions à valeurs réelles sur un espace mesurable . . . 720
13.4.4 Fonction étagée . . . 726
13.4.5 Fonctions réelle à variables réelles . . . 729
13.5 Intégrale par rapport à une mesure . . . 729
13.5.1 Quelque propriétés . . . 731
13.5.2 Permuter limite et intégrale . . . 734
13.5.2.1 Convergence uniforme . . . 734
13.5.2.2 Convergence monotone . . . 735
13.5.2.3 Convergence dominée de Lebesgue . . . 738
13.5.3 Produit d'une mesure par une fonction (mesure à densité) . . . 739
13.5.4 Mesure et topologie . . . 740
13.6 Propriétés . . . 741
13.7 Mesure à densité . . . 742
13.7.1 Théorème de Radon-Nikodym . . . 742
13.7.2 Mesure complexe . . . 744
13.7.3 Théorème d'approximation . . . 745
13.8 Tribu produit, mesure produit . . . 745
13.8.1 Produit d'espaces mesurables . . . 745
13.8.2 Le cas des boréliens . . . 746
13.8.3 Produit de mesures . . . 747
13.9 Mesure de Lebesgue sur $\mathbb{R}$ . . . 752
13.9.1 Mesure et tribu de Lebesgue . . . 756
13.9.2 Propriétés de la mesure de Lebesgue . . . 757
13.9.3 Ensemble de Cantor . . . 762
13.9.4 Ensemble de Vitali (non mesurable) . . . 764

13.10 Tribu et mesure de Lebesgue sur $\mathbb{R}^d$ . . . 765
13.10.1 Ensembles négligeables . . . 766
13.10.2 Parties et fonctions mesurables . . . 767
13.10.3 Propriétés d'unicité . . . 767
13.10.4 Régularité . . . 769
13.11 Propriétés de l'intégrale de Lebesgue . . . 769
13.11.1 Théorème de la moyenne . . . 769
13.11.2 Primitives et intégrales . . . 770
13.11.3 Exemples et applications . . . 772
13.11.4 Intégrales impropres . . . 773
13.12 Changement de variables dans une intégrale multiple . . . 775
13.12.1 Des lemmes . . . 775
13.12.2 Déterminant et mesure de Lebesgue . . . 776
13.12.3 Le théorème et sa démonstration . . . 777
13.12.4 Exemples . . . 782
13.13 Théorème de Fubini-Tonelli et de Fubini . . . 783

14 Intégration (Vol 2) 793
14.1 Intégrale sur une variété . . . 793
14.1.1 Mesure sur une carte . . . 793
14.1.1.1 Exemple : la mesure de la sphère . . . 794
14.1.2 Intégrale sur une carte . . . 794
14.1.3 Exemples . . . 795
14.1.4 Orientation . . . 795
14.1.5 Formes différentielles . . . 797
14.1.6 Intégrale d'une fonction sur une sous-variété . . . 798
14.2 Intégrales curvilignes . . . 799
14.2.1 Chemins de classe C^1 . . . 799
14.2.2 Intégrer une fonction . . . 799
14.2.3 Intégrer un champ de vecteurs . . . 801
14.2.4 Intégrer une forme différentielle sur un chemin . . . 801
14.2.5 Intégration d'une forme différentielle sur un chemin . . . 802
14.2.6 Interprétation physique : travail . . . 803
14.2.7 Intégrer un champs de vecteurs sur un bord en $2D$ . . . 804
14.2.8 Intégrer une forme différentielle sur un bord en $2D$ . . . 804
14.2.9 Intégrer une forme différentielle sur un bord en $3D$ . . . 804
14.2.10 Intégrer d'un champ de vecteurs sur un bord en $3D$ . . . 805
14.2.11 Dérivées croisées et forme différentielle exacte . . . 805
14.3 Surfaces paramétrées . . . 806
14.3.1 Graphe d'une fonction . . . 807
14.3.2 Intégrale sur une partie de $\mathbb{R}^m$ . . . 808
14.4 Intégrales de surface . . . 809
14.4.1 Intégrale d'un champ de vecteurs . . . 809
14.5 Intégrales de surface . . . 810
14.5.1 Aire d'une surface paramétrée . . . 810
14.5.2 Intégrale d'une fonction sur une surface . . . 811
14.5.3 Aire d'une surface de révolution . . . 812
14.5.4 Intégrale d'une 2-forme . . . 813
14.6 Flux d'un champ de vecteurs à travers une surface . . . 814
14.7 Divergence, Green, Stokes . . . 817
14.7.1 Théorème de la divergence . . . 817
14.7.2 Formule de Green . . . 818
14.7.3 Formule de Stokes . . . 820
14.7.3.1 Quelle est la bonne orientation ? . . . 820

14.8 Résumé des intégrales vues . . . 821
14.8.1 L'intégrale d'une fonction sur les réels . . . 821
14.8.2 Intégrale d'une fonction sur un chemin . . . 821
14.8.3 Intégrale d'une fonction sur une surface . . . 821
14.8.4 Intégrale d'une fonction sur un volume . . . 822
14.8.5 Conclusion pour les fonctions . . . 823
14.8.6 Circulation d'un champ de vecteurs . . . 823
14.8.7 Flux d'un champ de vecteurs . . . 823
14.8.8 Conclusion pour les champs de vecteurs . . . 823
14.8.9 Attention pour les surfaces fermées ! . . . 824
14.9 Intégrales convergeant uniformément . . . 825
14.9.1 Définition et propriété . . . 825
14.9.2 Critères de convergence uniforme . . . 826
14.10Fonctions définies par une intégrale . . . 827
14.10.1 Continuité sous l'intégrale . . . 827
14.10.2 Le coup du compact . . . 829
14.10.3 Dérivabilité sous l'intégrale . . . 829
14.10.4 Absolue continuité . . . 830
14.10.5 Différentiabilité sous l'intégrale . . . 832
14.11Formes différentielles exactes et fermées . . . 835
14.12Théorème d'Abel angulaire . . . 837
14.12.1 Passage à la limite sous le signe intégral . . . 841
14.12.2 Intégrale en dimension un . . . 841
14.12.3 Intégrales convergentes . . . 842
14.12.4 La méthode de Rothstein-Trager . . . 842
14.13Rappel sur les intégrales usuelles . . . 847
14.14Intégrales le long de chemins . . . 848
14.14.1 Circulation d'un champ de vecteur . . . 848
14.15Circulation d'un champ conservatif . . . 850
14.16Intégration de fonction à deux variables . . . 851
14.16.1 Intégration sur un domaine rectangulaire . . . 851
14.16.2 Intégration sur un domaine non rectangulaire . . . 852
14.16.3 Changement de variables . . . 855
14.16.3.1 Le cas des coordonnées polaires . . . 855
14.16.3.2 Les coordonnées cylindriques . . . 856
14.16.3.3 Coordonnées sphériques . . . 857
14.16.3.4 Un autre système utile . . . 857
14.17Les intégrales triples . . . 858
14.17.1 Volume . . . 860
14.18Un petit peu plus formel . . . 861
14.18.1 Intégration sur un domaine non rectangulaire . . . 861
14.18.2 Changement de variables . . . 862
14.18.2.1 Coordonnées polaires . . . 863
14.18.2.2 Coordonnées sphériques . . . 863
14.19Aire et primitive . . . 863
14.19.1 Longueur d'arc de courbe . . . 864
14.19.2 Aire de révolution . . . 864
14.20L'aire en dessous d'une courbe . . . 864
14.21Propriétés des intégrales . . . 866
14.22Techniques d'intégration . . . 867
14.22.1 Intégration par parties . . . 868
14.22.2 Changement de variables – pour trouver des primitives . . . 869
14.22.3 Changement de variables – pour calculer des intégrales . . . 870

14.22.4 Intégrations des fractions rationnelles réduites . . . 872
14.22.5 Quelques formules à connaître . . . 873
14.23 Trucs et astuces de calcul d'intégrales . . . 873
14.23.1 Quelques intégrales «usuelles» . . . 874
14.23.2 Reformer un carré au dénominateur . . . 876
14.23.3 Décomposer en fractions simples . . . 877
14.24 Constructions plus naïves de l'intégrale dans le cas réel . . . 877
14.24.1 Mesure de Lebesgue, version rapide . . . 877
14.24.2 Pavés et subdivisions . . . 878
14.24.3 Intégrale d'une fonction en escalier . . . 881
14.24.4 Intégrales partielles . . . 881
14.24.5 Réduction d'une intégrale multiple . . . 882
14.24.6 Propriétés de l'intégrale . . . 883
14.24.7 Intégrales multiples, cas général . . . 884
14.24.8 Réduction d'une intégrale multiple . . . 885
14.24.9 Intégrales sur des parties de $\mathbb{R}^2$ . . . 886
14.24.10 Intégrales sur des parties de $\mathbb{R}^3$ . . . 889
14.24.11 Fonctions et ensembles non bornés . . . 892

15 Suites et séries de fonctions (Vol 2) 893
15.1 Différentielle et dérivée complexe . . . 893
15.2 Série de fonctions . . . 896
15.2.1 Intégration de séries de fonctions . . . 899
15.2.2 Différentiabilité . . . 900
15.3 Séries entières . . . 901
15.3.1 Disque de convergence . . . 901
15.3.2 Propriétés de la somme . . . 904
15.3.3 Dérivation . . . 908
15.3.4 Intégration . . . 910
15.4 Dénombrement des solutions d'une équation diophantienne . . . 911
15.5 Exponentielle et logarithme . . . 912
15.6 Vitesses de x^α, de l'exponentielle et du logarithme . . . 919
15.7 Nombres de Bell . . . 922
15.8 Trigonométrie hyperbolique . . . 924
15.9 Séries entières de matrices . . . 925
15.9.1 Différentiabilité . . . 925
15.10 Étude d'asymptote . . . 927
15.11 Développement en série . . . 928
15.11.1 Série génératrice d'une suite . . . 928
15.11.2 Développement en série et Taylor . . . 929
15.11.3 Resommer une série . . . 931
15.11.3.1 Les sommes du type $\sum_n P(n)x^n$ . . . 931
15.11.3.2 Les sommes du type $\sum_n x^n/P(n)$ . . . 932
15.11.3.3 Sage, primitives et logarithme complexe . . . 934
15.11.3.4 Nombres de Bell . . . 936
15.12 Séries entières de matrices . . . 936
15.12.1 Rayon de convergence . . . 936
15.12.2 Convergence et rayon spectral . . . 936
15.12.3 Exponentielle et logarithme de matrice . . . 938
15.13 Lemme de Borel . . . 941
15.13.1 Fonctions plateaux . . . 941
15.13.2 Le lemme de Borel . . . 942

16 Trigonométrie, isométries (Vol 3) 945
16.1 Isométries de l'espace euclidien . . . 945
16.1.1 Structure du groupe Isom($\mathbb{R}^n$) . . . 945
16.1.2 Classification des isométries de $\mathbb{R}$ . . . 949
16.1.3 Segment, plan médiateur et équidistance . . . 950
16.1.4 Isométries du tétraèdre régulier . . . 950
16.2 Trigonométrie . . . 951
16.2.1 Définitions, périodicité et quelque valeurs remarquables . . . 951
16.2.2 Exemples . . . 957
16.2.3 Très modeste approximation de π . . . 959
16.2.4 Cercle trigonométrique . . . 959
16.2.5 Les fonctions tangente et arc tangente . . . 960
16.2.6 La fonction arc sinus . . . 962
16.2.7 La fonction arc cosinus . . . 964
16.2.8 Une meilleure approximation de π . . . 965
16.2.9 Forme trigonométrique des nombres complexes . . . 966
16.2.10 Angle entre deux vecteurs . . . 967
16.2.11 Aire du parallélogramme . . . 968
16.3 Classification des isométries dans $\mathbb{R}^2$ . . . 969
16.3.1 Réflexions . . . 969
16.3.2 Rotations . . . 971
16.3.3 Rotations vectorielles . . . 975
16.3.4 Angle . . . 976
16.3.5 Classification . . . 978
16.4 Isométries dans $\mathbb{R}^n$ . . . 981
16.4.1 Groupes finis d'isomorphismes . . . 985
16.4.2 Théorème de Sylvester . . . 987
16.4.3 Groupe diédral . . . 988
16.4.3.1 Définition et générateurs : vue géométrique . . . 988
16.4.3.2 Générateurs : vue abstraite . . . 990
16.4.3.3 Classes de conjugaison . . . 993
16.4.3.4 Le compte pour n pair . . . 994
16.4.3.5 Le compte pour n impair . . . 994
16.4.4 Applications : du dénombrement . . . 994
16.4.4.1 Le jeu de la roulette . . . 994
16.4.4.2 L'affaire du collier . . . 995
16.5 Angles et rotations dans $\mathbb{R}^2$ . . . 995
16.5.1 Angles orientés, rotations . . . 995
16.5.2 Angles et nombres complexes . . . 1000
16.5.3 Sous-groupes finis de SO(2) . . . 1001
16.6 Un peu de structure de O(n) . . . 1001
16.6.1 Valeurs propres dans O(n) . . . 1001
16.6.2 Sous-groupes finis de SO(3) . . . 1004
16.7 Coordonnées polaires . . . 1013
16.7.1 Les coordonnées polaires . . . 1013
16.7.1.1 Transformation inverse : théorie . . . 1013
16.7.1.2 Transformation inverse : pratique . . . 1014
16.7.2 Coordonnées polaires comme changement de variables pour intégrales . . . 1015
16.7.3 Changement de variables . . . 1017
16.7.3.1 Coordonnées polaires . . . 1017
16.7.3.2 Coordonnées sphériques . . . 1018
16.7.4 Coordonnées polaires . . . 1019
16.7.5 Coordonnées cylindriques . . . 1019

16.7.6 Coordonnées sphériques . . . 1020
16.7.7 Coordonnées cylindriques et sphériques . . . 1020
16.7.8 Récapitulatif des changements de variables . . . 1021
16.7.8.1 Coordonnées polaires . . . 1021
16.7.8.2 Coordonnées cylindriques . . . 1021
16.7.8.3 Coordonnées sphériques . . . 1021
16.7.9 une limite . . . 1022

17 Représentations et caractères (Vol 3) 1023
17.1 Représentations et caractères . . . 1023
17.1.1 Crochet de dualité et transformée de Fourier . . . 1025
17.1.2 Groupes non abéliens . . . 1026
17.1.3 Représentations linéaires des groupes finis . . . 1026
17.1.4 Module . . . 1027
17.1.5 Structure hermitienne . . . 1029
17.1.6 Caractères . . . 1029
17.2 Équivalence de représentations et caractères . . . 1030
17.2.1 Représentation régulière . . . 1032
17.2.2 Caractères et représentations : suite et fin . . . 1033
17.3 Représentation produit tensoriel . . . 1036
17.4 Exemple sur le groupe symétrique . . . 1036
17.5 Table des caractères du groupe symétrique S_4 . . . 1037
17.5.1 Calculs à partir de rien ou presque . . . 1037
17.5.2 Représentation de S_4 via les isométries du tétraèdre . . . 1039
17.5.3 À propos de la représentation ρ_u . . . 1040
17.6 Table de caractères du groupe diédral . . . 1041
17.6.1 Représentations de dimension un . . . 1041
17.6.2 Représentations de dimension deux . . . 1042
17.6.3 Le compte pour n pair . . . 1044
17.6.4 Le compte pour n impair . . . 1044

18 Suite de l'analyse (Vol 3) 1045
18.1 Densité des polynômes . . . 1045
18.1.1 Théorème de Stone-Weierstrass . . . 1045
18.2 Primitive de fonction continue . . . 1048
18.2.1 Théorème taubérien de Hardi-Littlewood . . . 1049
18.2.2 Théorème de Müntz . . . 1053
18.3 Théorèmes de point fixe . . . 1056
18.3.1 Points fixes attractifs et répulsifs . . . 1056
18.3.2 Picard . . . 1056
18.3.3 Brouwer . . . 1059
18.3.4 Théorème de Schauder . . . 1061
18.3.5 Théorème de Markov-Kakutani et mesure de Haar . . . 1062
18.4 Théorèmes de point fixes et équations différentielles . . . 1064
18.4.1 Théorème de Cauchy-Lipschitz . . . 1064
18.4.2 Théorème de Cauchy-Arzella . . . 1071
18.5 Théorèmes d'inversion locale et de la fonction implicite . . . 1072
18.5.1 Mise en situation . . . 1072
18.5.2 Théorème d'inversion locale . . . 1073
18.5.3 Théorème de la fonction implicite . . . 1076
18.5.4 Exemple . . . 1078
18.6 Décomposition polaire (régularité) . . . 1079
18.7 Théorème de Von Neumann . . . 1081
18.8 Recherche d'extrema . . . 1084

18.8.1 Extrema à une variable . . . 1084
18.8.2 Extrema libre . . . 1085
18.8.3 Extrema et Hessienne . . . 1086
18.8.4 Un peu de recettes de cuisine . . . 1088
18.8.5 Extrema liés . . . 1088
18.9 Fonctions convexes . . . 1090
18.9.1 Inégalité des pentes . . . 1091
18.9.2 Convexité et régularité . . . 1092
18.9.3 Dérivées d'une fonction convexe . . . 1092
18.9.4 Graphe d'une fonction convexe . . . 1094
18.9.5 Convexité et hessienne . . . 1098
18.9.6 Quelque inégalités . . . 1102
18.9.6.1 Inégalité de Jensen . . . 1102
18.9.6.2 Inégalité arithmético-géométrique . . . 1102
18.9.6.3 Inégalité de Kantorovitch . . . 1103
18.10 Algorithme du gradient à pas optimal . . . 1104
18.11 Ellipsoïde de John-Loewer . . . 1108
18.12 Formes quadratiques, signature, et lemme de Morse . . . 1111
18.12.1 Lemme de Morse . . . 1115
18.13 Variétés . . . 1117
18.13.1 Introduction . . . 1117
18.13.2 Définition et propriétés . . . 1117
18.13.3 Espace tangent . . . 1118
18.14 Prolongement de fonctions et complétion d'espaces métriques . . . 1119
18.15 Un petit extra . . . 1125

19 Arc paramétré (Vol 3) 1127
19.1 Définitions . . . 1127
19.2 Longueur d'arc . . . 1127
19.3 Abscisse curviligne . . . 1130
19.3.1 Formule intégrale de la longueur . . . 1131
19.4 Suite du chapitre . . . 1135
19.5 Courbes paramétrés . . . 1136
19.5.1 Définitions et exemples . . . 1136
19.6 Élément de longueur . . . 1138
19.6.1 Élément de longueur : cartésiennes . . . 1138
19.6.2 Élément de longueur : polaires (1) . . . 1138
19.6.3 Élément de longueur : polaires (2) . . . 1138
19.6.4 Approximation de la longueur par des cordes . . . 1140
19.7 Arc géométrique . . . 1142
19.7.1 Abscisse curviligne et paramétrisation normale . . . 1143
19.7.2 Tangente à une courbe paramétrée . . . 1148
19.8 Repère de Frenet . . . 1150
19.8.1 Torsion . . . 1152
19.9 Hors des coordonnées normales . . . 1153
19.10 Tracer des courbes paramétriques dans $\mathbb{R}^2$ . . . 1156
19.11 Courbes planes . . . 1157
19.11.1 Angle . . . 1157
19.11.2 Courbure signée . . . 1158
19.11.3 Degré, indice et homotopie . . . 1161
19.12 Courbes fermées planes . . . 1167
19.12.1 Cercle circonscrit . . . 1167
19.12.2 Description locale . . . 1169
19.12.3 Enveloppe convexe . . . 1170

19.12.4 Courbure et convexité . . . 1173
19.12.5 Théorème des quatre sommets . . . 1174
19.12.6 Le théorème de Jordan . . . 1176

20 Géométrie hyperbolique (Vol 3) 1177
20.1 Inversion . . . 1177
20.1.1 Cercles perpendiculaires . . . 1177
20.1.2 Inversion . . . 1178

21 Espaces projectifs (Vol 3) 1183
21.1 Sous espaces projectifs . . . 1183
21.2 Espace projectifs comme «complétés» d'espaces affines . . . 1185
21.3 Théorème de Pappus . . . 1188
21.4 Homographies . . . 1189
21.4.1 Homographies . . . 1189
21.4.2 Le groupe projectif . . . 1190
21.4.3 Repères projectifs . . . 1191
21.4.4 Identifications $P(\mathbb{K}^2)$ vers $\mathbb{K} \cup \{\infty\}$ . . . 1195
21.4.5 Birapport . . . 1196
21.5 Coordonnées homogènes . . . 1202
21.5.1 Dualité . . . 1202
21.5.2 Polynômes . . . 1203
21.6 La sphère de Riemann $P_1(\mathbb{C})$ . . . 1205
21.6.1 Éléments de géométrie dans $P_1(\mathbb{C})$ . . . 1205
21.6.1.1 Équation complexe d'une droite . . . 1206
21.6.1.2 Équation complexe d'un cercle . . . 1206
21.6.1.3 Cercle-droite . . . 1207
21.6.1.4 Rotation-homothétie . . . 1208
21.6.1.5 Application linéaire . . . 1208
21.6.1.6 Inversion . . . 1208
21.6.2 Homographies . . . 1209
21.6.3 Birapport . . . 1214
21.6.4 Division harmonique . . . 1217
21.6.5 Groupe circulaire . . . 1221
21.6.6 Action du groupe modulaire . . . 1224

22 Analyse vectorielle (Vol 3) 1229
22.1 Le théorème de Green . . . 1229
22.2 Théorème de la divergence dans le plan . . . 1233
22.2.1 La convention de sens de parcours . . . 1233
22.2.2 Théorème de la divergence . . . 1234
22.3 Théorème de Stokes . . . 1234
22.4 Théorème de Gauss . . . 1236
22.5 Coordonnées curvilignes . . . 1238
22.5.1 Base locale . . . 1238
22.5.2 Importance de l'orthogonalité . . . 1238
22.5.3 Coordonnées polaires . . . 1240
22.5.4 Coordonnées cylindriques . . . 1240
22.5.5 Coordonnées sphériques . . . 1241
22.5.6 Gradient en coordonnées curvilignes . . . 1241
22.5.6.1 Coordonnées sphériques . . . 1242
22.5.7 Divergence en coordonnées curvilignes . . . 1242
22.5.7.1 Coordonnées cylindriques . . . 1244
22.5.7.2 Coordonnées sphériques . . . 1244

22.5.8 Laplacien en coordonnées curvilignes orthogonales . . . 1244
22.5.9 Rotationnel en coordonnées curvilignes orthogonales . . . 1245
22.5.9.1 Coordonnées cylindriques . . . 1246
22.5.9.2 Coordonnées sphériques . . . 1246
22.6 Les formules . . . 1246
22.6.1 Coordonnées polaires . . . 1246
22.6.2 Coordonnées cylindriques . . . 1246
22.6.3 Coordonnées sphériques . . . 1247

23 Espaces de Hilbert (Vol 3) 1249
23.1 Espaces de Hilbert . . . 1249
23.1.1 Sous-espace vectoriel fermé ? ? ? . . . 1250
23.2 Théorème de la projection . . . 1250
23.3 Systèmes orthogonaux et bases . . . 1253
23.3.1 Orthogonal d'une partie . . . 1254
23.3.2 Dual, théorème de représentation de Riesz . . . 1255
23.3.3 Séparabilité . . . 1256
23.3.4 Base hilbertienne . . . 1258
23.3.5 Décomposition dans une base hilbertienne . . . 1263
23.3.6 Digression sur les normes opérateurs . . . 1267
23.3.7 Applications linéaires et continuité . . . 1268
23.4 Théorème de Kochen-Specker . . . 1269
23.5 Théorème de Lax-Milgram . . . 1270

24 Analyse fonctionnelle (Vol 3) 1275
24.1 Théorème d'Ascoli . . . 1275
24.2 Théorème de Banach-Steinhaus . . . 1276
24.3 Espaces L^p . . . 1278
24.3.1 Généralités . . . 1278
24.3.2 L'espace L^∞ . . . 1281
24.3.3 Quelque identifications . . . 1281
24.3.4 Inégalité de Jensen, Hölder et de Minkowski . . . 1282
24.3.5 Ni inclusions ni inégalités . . . 1284
24.3.6 Complétude . . . 1286
24.3.7 Densité des fonctions infiniment dérivables à support compact . . . 1290
24.3.8 Approximation . . . 1293
24.4 Convolution . . . 1294
24.4.1 Approximation de l'unité . . . 1295
24.4.2 Densité des polynôme trigonométrique . . . 1297
24.5 L'espace L^2 . . . 1299
24.5.1 Coefficients et série de Fourier . . . 1300
24.5.2 Dualité et théorème de représentation de Riesz . . . 1302
24.6 Théorèmes de Hahn-Banach . . . 1310
24.7 Théorème de Tietze . . . 1312
24.8 Espace de Schwartz . . . 1315
24.8.1 Topologie . . . 1316
24.8.2 Produit de convolution . . . 1319

25 Analyse complexe (Vol 3) 1321
25.1 Fonctions holomorphes . . . 1321
25.1.1 Équations de Cauchy-Riemann . . . 1321
25.1.2 Intégrales sur des chemins fermés . . . 1322
25.1.3 Lacets, indice et homotopie . . . 1325
25.1.4 Théorème de Cauchy et analycité . . . 1326

25.1.5 Théorème de Brouwer en dimension 2 . . . 1329
25.1.6 Principe des zéros isolés . . . 1330
25.1.7 Prolongement de fonctions holomorphes . . . 1331
25.1.8 Théorème de Runge . . . 1331
25.2 Intégrales de fonctions holomorphes . . . 1333
25.2.1 Mesure de Radon . . . 1338
25.3 Conditions équivalentes à l'holomorphie . . . 1339
25.4 Singularités, pôles et méromorphe . . . 1339
25.5 Fonctions d'Euler . . . 1340
25.5.1 Euler et factorielle . . . 1343
25.6 Partition d'un entier en parts fixées . . . 1344
25.7 Exponentielle et logarithme complexe . . . 1347
25.7.1 Définition et propriétés de l'exponentielle . . . 1347
25.7.2 Intégrale de Fresnel . . . 1349
25.7.3 Logarithme complexe . . . 1351
25.7.3.1 La fonction argument . . . 1351
25.7.3.2 Une définition possible du logarithme . . . 1354
25.7.3.3 Pas plus de continuité . . . 1356
25.7.3.4 Pas d'unicité : autres déterminations de l'argument . . . 1357
25.7.3.5 Pas d'unicité : développement en série . . . 1358
25.7.3.6 Pas d'unicité : laquelle choisir ? . . . 1358
25.7.3.7 Logarithme comme primitive . . . 1359
25.8 Théorème de Weierstrass . . . 1360
25.9 Théorème de Montel . . . 1361
25.10Espaces de Bergman . . . 1363

26 Série de Fourier (Vol 3) 1367
26.1 Densité des polynômes trigonométriques . . . 1367
26.1.1 Convergence pour les fonctions continues (via Weierstrass) . . . 1367
26.1.2 Convergence pour les fonctions continues (via Fejér) . . . 1367
26.1.3 Densité dans L^p . . . 1370
26.1.4 Suite équirépartie, critère de Weyl . . . 1371
26.2 Fonctions de Dirichlet . . . 1373
26.3 Coefficients et série de Fourier . . . 1374
26.3.1 Le contre-exemple que nous attendions tous . . . 1377
26.3.2 Inégalité isopérimétrique . . . 1379
26.3.3 À propos des coefficients . . . 1380

27 Transformation de Fourier (Vol 4) 1383
27.1 Transformée de Fourier dans $L^1(\mathbb{R}^d)$ . . . 1383
27.1.1 Formule sommatoire de Poisson . . . 1386
27.2 Transformée de Fourier dans l'espace de Schwartz . . . 1388
27.2.1 Quelque transformées de Fourier . . . 1391
27.3 Suite régularisante . . . 1393
27.3.1 Formule d'inversion . . . 1395
27.4 Transformée de Fourier sur $L^2(\mathbb{R}^d)$ . . . 1397
27.4.1 Extension de $L^1 \cap L^2$ vers L^2 . . . 1397

28 Distributions (Vol 4) 1401
28.1 Topologie . . . 1401
28.1.1 Dérivée partielle au sens faible . . . 1401
28.2 Distributions . . . 1404
28.2.1 Multiplication d'une distribution par une fonction . . . 1406
28.2.2 Dérivée de distribution . . . 1406

28.2.3 Ordre et support d'une distribution . . . 1407
28.3 Distributions tempérées . . . 1410
28.3.1 Topologie . . . 1411
28.3.2 Distributions associées à des fonctions . . . 1411
28.3.3 Composition avec une fonction . . . 1412
28.3.4 Transformée de Fourier d'une distribution tempérée . . . 1412
28.3.5 Convolution d'une distribution par une fonction . . . 1412
28.3.6 Approximation de la distribution de Dirac . . . 1413
28.3.7 Peigne de Dirac . . . 1416
28.4 L'espace $C^\infty(\mathbb{R}, \mathscr{D}'(\mathbb{R}^d))$ . . . 1417
28.4.1 Dérivation . . . 1419
28.5 L'espace $C^\infty(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d))$ . . . 1420
28.5.1 Propriétés générales . . . 1420
28.5.2 Dérivation . . . 1422
28.6 Une équation de distribution . . . 1424

29 Espaces de Sobolev, équations elliptiques (Vol 4) 1427
29.1 Espaces de Sobolev . . . 1427
29.1.1 Sur un intervalle de $\mathbb{R}$ . . . 1427
29.1.2 Sur un ouvert de $\mathbb{R}^n$ . . . 1433
29.1.2.1 Définition . . . 1433
29.1.3 Espace de Sobolev fractionnaire . . . 1434
29.2 Trace . . . 1437
29.3 Théorème de plongement . . . 1440

30 Équations différentielles ordinaires (Vol 4) 1445
30.1 Équation homogène, solution particulière . . . 1446
30.2 Que faire avec $f(z)dz = g(t)dt$? . . . 1446
30.3 Équations linéaires du premier ordre . . . 1448
30.3.1 Pourquoi la variation des constantes fonctionne toujours ? . . . 1449
30.4 Équations à variables séparées . . . 1450
30.4.1 La méthode rapide . . . 1451
30.4.2 La méthode plus propre . . . 1451
30.4.3 Les théorèmes . . . 1451
30.5 Équations linéaires d'ordre supérieur . . . 1453
30.5.1 Équations et systèmes linéaire à coefficients constants . . . 1453
30.5.2 Si les coefficients ne sont pas constants ? . . . 1454
30.6 Système d'équations linéaires . . . 1454
30.6.1 La magie de l'exponentielle... . . . 1454
30.6.2 ...mais la difficulté . . . 1455
30.6.3 La recette . . . 1455
30.6.4 Système d'équations linéaires avec matrice constante . . . 1456
30.6.5 Système d'équations linéaires avec matrice non constante . . . 1456
30.7 Réduction de l'ordre . . . 1456
30.8 Autour de Cauchy-Lipschitz . . . 1458
30.8.1 Fuite des compacts et explosion en temps fini . . . 1458
30.8.2 Écart entre deux conditions initiales . . . 1460
30.8.3 Flot d'un champ de vecteurs . . . 1461
30.8.4 Stabilité de Lyapunov . . . 1475
30.8.5 Système proie et prédateurs : Lokta-Voltera . . . 1479
30.9 Équation du second ordre . . . 1482
30.9.1 Wronskien . . . 1482
30.9.2 Avec second membre . . . 1482
30.9.3 Équation $y'' + q(t)y = 0$ . . . 1483

30.9.4 Équation de Hill 1484
30.10 Différents types d'équations différentielles 1487
30.10.1 Équation homogène 1487
30.10.2 Équation de Bernoulli 1488
30.10.3 Équation de Riccati 1488
30.10.4 Équation différentielle exacte 1488
30.10.4.1 Résolution lorsque tout va bien 1488
30.10.4.2 Facteur intégrant (quand tout ne va pas bien) 1489
30.11 Distributions pour les équations différentielles 1489
30.11.1 Équation de Schrödinger 1490
30.12 Équations différentielles du premier ordre 1493
30.13 Équations différentielles du premier ordre à variables séparables 1496
30.14 Équations différentielles linéaires du premier ordre 1498
30.14.1 Méthode de variation de la constante 1500
30.15 Équations différentielles linéaires du second ordre 1501
30.15.1 Résolution des équations différentielles linéaires du second ordre homogènes à coefficients constants 1501
30.15.2 Équations différentielles linéaires du second ordre à coefficients constants, non homogènes 1503
30.16 Fonction de Green 1504

31 Équations aux dérivées partielles (Vol 4) 1507
31.1 Symbole principal, équation des caractéristiques 1507
31.2 Méthode des caractéristiques pour l'ordre 1 1507
31.2.1 Un exemple complet un peu minimal 1508
31.2.2 Un théorème d'existence et d'unicité 1510
31.3 Méthode des caractéristique pour l'ordre 2 1513
31.3.1 Principe général 1513
31.3.2 Exemple : l'équation d'onde 1514
31.4 Classification des équations du second ordre 1516
31.4.1 Problème au limite 1517
31.5 Principe du maximum 1518
31.6 Quelque exemples 1522
31.6.1 Un changement de variables 1522

32 Numérique (Vol 4) 1525
32.1 Introduction 1525
32.2 Représentations numériques 1525
32.2.1 Entier relatif en complément à deux (binaire) 1525
32.2.2 Représentation en virgule flottante 1527
32.2.3 Simple précision, IEEE-754 1527
32.3 Problèmes pour écrire des nombres 1530
32.3.1 Troncature : la base 1530
32.3.2 Troncature : le drift 1531
32.3.3 Quelque bonnes règles 1532
32.3.4 Erreur de "cancellation" 1532
32.3.5 Calcul d'une dérivée 1534
32.3.6 Erreur d'absorption 1535
32.4 Conditionnement et stabilité 1535
32.4.1 Comment choisir et penser le K ? 1538
32.5 Un peu de points fixes 1539
32.5.1 Choix de la fonction à point fixe 1539
32.5.2 Convergence quadratique 1541
32.5.3 Convergence 1542

32.6 Méthode de Newton . . . 1543
32.6.1 «Justification» par la formule par Taylor . . . 1544
32.6.2 «Justification» par points fixes . . . 1544
32.6.3 Convergence de la méthode de Newton . . . 1545
32.6.4 Formalisation de l'algorithme . . . 1547
32.6.5 Caractéristiques . . . 1548
32.6.6 Exemple de la racine carré . . . 1548
32.6.7 Si multiplicité . . . 1549
32.6.8 Et la dérivée ? . . . 1549
32.6.9 Méthode de Newton : le cas général . . . 1549
32.7 Estimation de l'ordre de convergence . . . 1551
32.8 Autres méthodes . . . 1552
32.8.1 Méthode de Schröder . . . 1552
32.8.2 Halley . . . 1552
32.9 Méthode des sécantes variables . . . 1553
32.9.1 Aitken . . . 1553
32.10Équations algébrique . . . 1554
32.10.1 Résoudre un système linéaire . . . 1554
32.10.2 Caractéristiques . . . 1555
32.10.3 Définitions . . . 1555
32.11Équations non linéaire . . . 1556
32.11.1 Méthode de bisection . . . 1557
32.12Efficacité . . . 1559
32.13Exercices . . . 1560
32.14Approximations de fonctions . . . 1563
32.14.1 Critère d'interpolation . . . 1563
32.14.2 Base de Newton . . . 1564
32.14.3 Méthode des minima quadratiques . . . 1566
32.14.4 Notre espace de Hilbert . . . 1567
32.14.5 Droite de régression . . . 1568
32.15Conditionnement d'une matrice . . . 1569
32.15.1 Perturbation du vecteur . . . 1571
32.15.2 Perturbation de la matrice . . . 1572
32.16Système linéaires (généralités) . . . 1573
32.16.1 Les méthodes directes . . . 1574
32.16.2 Méthodes itératives . . . 1575
32.17Système linéaires (méthodes directes) . . . 1575
32.17.1 Inversion de matrice triangulaire . . . 1575
32.17.2 Transformation gaussienne . . . 1575
32.17.3 Méthode de Gauss pour résoudre des systèmes d'équations linéaires . . . 1577
32.17.4 Méthode de Gauss sans pivot (décomposition LU) . . . 1578
32.17.5 Matrice de permutation élémentaires . . . 1581
32.18Méthode de Gauss avec pivot partiel (décomposition PLU) . . . 1582
32.18.1 L'idée . . . 1582
32.18.2 Le théorème . . . 1583
32.18.3 D'un point de vue algorithmique . . . 1587
32.18.4 Exemples . . . 1589
32.19Résolution de systèmes linéaires (suite) . . . 1592
32.19.1 Déterminant . . . 1592
32.19.2 Plusieurs termes indépendants . . . 1592
32.19.3 Cholesky . . . 1593
32.20Système linéaire (méthodes itératives) . . . 1596
32.20.1 La méthode générale . . . 1597

32.20.2 Jacobi . . . 1597
32.20.3 Gauss-Seidel . . . 1597
32.20.4 Autres . . . 1597
32.21 Indices connectés, matrice irréductible . . . 1597
32.22 Localisation des valeurs propres . . . 1599
32.22.1 Matrices à diagonale dominante . . . 1602
32.22.2 M-matrice . . . 1604

33 Méthodes des différences finies (Vol 4) 1609
33.1 Problèmes de dimension un . . . 1609
33.1.1 Un schéma à cinq points . . . 1610
33.1.1.1 Poser le système . . . 1610
33.1.1.2 Propriétés du système . . . 1612
33.1.2 Exemple . . . 1614
33.2 Problèmes de dimension deux . . . 1615
33.2.1 Discrétisation en croix . . . 1615
33.2.2 Discrétisation en carré . . . 1616
33.2.3 Résolution de la discrétisation en croix . . . 1617
33.3 Consistance, convergence . . . 1619
33.3.1 Définitions, mise en place . . . 1619
33.3.2 Exemple . . . 1621
33.3.3 Consistance, stabilité et convergence . . . 1622
33.3.4 Exemple : schéma à cinq points, laplacien en croix . . . 1623
33.4 Autres laplaciens . . . 1624
33.4.1 Travail avec le laplacien à 9 points . . . 1627

34 Variables aléatoires et théorie des probabilités (Vol 4) 1629
34.1 Espace de probabilité . . . 1629
34.2 Variables aléatoires . . . 1630
34.2.1 Indépendance . . . 1630
34.2.2 Lois conjointes et indépendance . . . 1633
34.2.3 Somme et produit de variables aléatoires indépendantes . . . 1634
34.2.4 Espérance . . . 1636
34.2.5 Variance . . . 1637
34.2.6 Covariance . . . 1638
34.2.7 Probabilité conditionnelle : événements . . . 1639
34.2.8 Espérance conditionnelle . . . 1640
34.2.9 Probabilité conditionnelle : tribu . . . 1646
34.2.10 Variables de Rademacher indépendantes . . . 1648
34.2.11 Un petit paradoxe . . . 1650
34.2.11.1 «Bonjour, je suis l'aînée» . . . 1650
34.2.11.2 «Bonjour» . . . 1651
34.2.11.3 Le parent qui répond aux questions . . . 1653
34.2.11.4 À propos des simulations . . . 1655
34.2.12 Inégalité de Jensen . . . 1655
34.2.13 Fonction de répartition . . . 1656
34.2.14 Fonction caractéristique . . . 1656
34.2.15 Fonction génératrice des moments, transformée de Laplace . . . 1657
34.2.16 Loi d'une variable aléatoire . . . 1658
34.2.17 Changement de variables . . . 1659
34.3 Convergence . . . 1660
34.4 Loi des grands nombres, théorème central limite . . . 1665
34.4.1 Loi des grands nombres . . . 1665
34.4.2 Théorème central limite . . . 1667

34.4.3 Marche aléatoire . . . 1670
34.5 Les lois usuelles . . . 1671
34.5.1 Loi de Bernoulli . . . 1671
34.5.2 Loi binomiale . . . 1672
34.5.3 Loi multinomiale . . . 1673
34.5.4 Loi géométrique . . . 1673
34.5.5 Loi de Poisson . . . 1674
34.5.6 Loi exponentielle . . . 1674
34.5.7 Approximation de la binomiale par une Poisson . . . 1677
34.5.8 Loi de Poisson et loi exponentielle . . . 1678
34.5.9 Loi normale . . . 1679
34.5.10 Vecteurs gaussiens . . . 1681
34.5.11 Variable aléatoire de Rademacher . . . 1685
34.5.12 Loi de Student . . . 1687
34.5.13 Indépendance, covariance et variance de somme . . . 1688
34.6 Estimation des grands écarts . . . 1688
34.7 Simulations de réalisations de variables aléatoires . . . 1692
34.7.1 Générateur uniforme . . . 1692
34.7.1.1 Première méthode . . . 1692
34.7.1.2 Seconde méthode . . . 1692
34.7.2 Simulation par inversion . . . 1692
34.7.2.1 Loi exponentielle . . . 1693
34.7.3 Algorithme de Box-Muller . . . 1693
34.7.4 Méthode du rejet . . . 1694
34.7.5 Simuler une loi géométrique à l'ordinateur . . . 1696
34.7.6 Simuler une loi exponentielle à l'ordinateur . . . 1696
34.7.7 Simuler une loi de Poisson à l'ordinateur . . . 1696
34.8 Sage . . . 1697
34.8.1 Loi exponentielle . . . 1697
34.8.2 Inverser des lois . . . 1697
34.9 Monte-Carlo . . . 1697
34.9.1 Intervalle de confiance . . . 1698
34.9.1.1 Principe . . . 1698
34.9.1.2 Échantillonnage préférentiel . . . 1699
34.9.1.3 Méthode de la variable de contrôle . . . 1700
34.9.1.4 Variables antithétiques . . . 1700
34.10 Résultats qui se démontrent avec des variables aléatoires . . . 1700
34.10.1 Nombres normaux . . . 1700
34.10.2 Théorème de Bernstein . . . 1702

35 Statistiques (Vol 4) 1707
35.1 Notations et hypothèses . . . 1707
35.2 Modèle statistique . . . 1707
35.3 Modèles d'échantillonnages . . . 1710
35.4 Estimation ponctuelle . . . 1713
35.5 Statistiques et estimateurs . . . 1715
35.5.1 Qualité des estimateurs . . . 1715
35.5.2 Méthode des moments . . . 1716
35.5.3 Méthode de substitution . . . 1718
35.5.4 Méthode du maximum de vraisemblance . . . 1718
35.5.5 Estimation d'une fonction de répartition . . . 1720
35.5.6 Espérance et variance d'un estimateur . . . 1723
35.6 Estimation par intervalle de confiance . . . 1723
35.6.1 Région de confiance . . . 1726

35.6.2 Fonction pivotale . . . 1726
35.6.3 Sondage de proportion . . . 1729
35.7 Estimer une densité lorsqu'on ne sait rien . . . 1730
35.7.1 Distance entre des mesures . . . 1731
35.7.2 Estimateur par fenêtres glissantes . . . 1731
35.8 Test d'hypothèses, prise de décision . . . 1734
35.8.1 Exemple : qualité des pièces d'usine . . . 1734
35.8.2 Exemple : la résistance d'un fil . . . 1734
35.8.3 Vocabulaire et théorie . . . 1735
35.8.4 Risque de première et seconde espèce . . . 1736
35.8.5 Modèle paramétrique de loi gaussienne . . . 1737
35.9 Tests paramétriques . . . 1738
35.10Tests d'adéquation . . . 1739

36 Chaînes de Markov à temps discret (Vol 4) 1745
36.1 Généralités . . . 1745
36.2 Chaînes de Markov sur un ensemble fini . . . 1746
36.3 Marche aléatoire sur $\mathbb{Z}$ . . . 1748
36.3.1 Chaînes de Markov homogènes . . . 1750
36.3.2 Graphe de transition . . . 1751
36.3.3 Chaîne de Markov définie par récurrence . . . 1751
36.3.3.1 Le cas général . . . 1751
36.3.3.2 Exemple : la file de réparation de machines à laver . . . 1753
36.4 Classification des états . . . 1753
36.4.1 Chaînes irréductibles . . . 1756
36.4.2 Nombre de visites . . . 1757
36.5 Mesure invariante . . . 1762
36.6 Convergence vers l'équilibre . . . 1764
36.7 Processus de Galton-Watson . . . 1766

37 Martingales (Vol 4) 1771
37.1 Convergence de martingales . . . 1771
37.2 Temps d'arrêt et martingale terminée . . . 1774
37.3 Décomposition de martingales . . . 1776
37.4 Problème de la ruine du joueur . . . 1778
37.4.1 Le cas où la pièce est truquée . . . 1779
37.4.1.1 Introduction d'une martingale . . . 1779
37.4.1.2 Finitude du temps d'arrêt . . . 1780
37.4.1.3 Temps moyen de jeu . . . 1780
37.4.1.4 Probabilité de victoire du joueur . . . 1781
37.4.2 Le cas où la pièce est non truquée . . . 1782
37.4.2.1 Probabilité de gagner . . . 1782
37.4.2.2 Temps moyen de jeu . . . 1784
37.4.3 Un petit complément . . . 1784

38 Processus de Poisson (Vol 4) 1785
38.1 Processus de Poisson . . . 1785
38.2 Quelques trucs sur la simulation . . . 1788
38.2.1 Le théorème central limite pour Markov . . . 1789
38.2.2 Feuille 5 . . . 1789
38.2.3 Feuille 6 . . . 1789
38.2.4 Feuille 7 . . . 1790
38.2.5 Simuler des lois conditionnelles . . . 1790

39 Utilisation dans les autres sciences (Vol 4) 1791
39.1 Démystification du MRUA . . . 1791
39.1.1 Preuve de la formule . . . 1791
39.1.2 Interprétation graphique . . . 1792
39.2 Relativité en mécanique newtonienne . . . 1792
39.2.1 Relativité du mouvement . . . 1792
39.2.2 Bob et Alice . . . 1792
39.3 Invariance de la vitesse de la lumière . . . 1793
39.3.1 Champ de gravitation et électrique . . . 1793
39.3.1.1 Finitude de la vitesse de propagation de la force électrique . . . 1793
39.3.1.2 Pourquoi pas la gravitation ? . . . 1793
39.3.2 Support du champ : pas d'éther . . . 1794
39.3.3 Le problème . . . 1794
39.4 Conséquences . . . 1794
39.4.1 Ligne d'univers . . . 1794
39.4.2 Transformations de Lorentz . . . 1795
39.4.3 Conditions d'existence . . . 1798
39.4.4 La notion d'intervalle . . . 1799
39.4.4.1 En mécanique newtonienne . . . 1799
39.4.4.2 En mécanique relativiste . . . 1799
39.4.5 Le cône de lumière d'un point . . . 1799
39.4.6 Contraction des longueurs . . . 1800
39.4.7 Dilatation des intervalles de temps . . . 1800
39.4.8 Invariance de l'intervalle . . . 1801
39.4.8.1 Rappel de trigonométrie hyperbolique . . . 1802
39.4.8.2 Les transformations de Lorentz (bis) . . . 1804
39.4.9 Vitesse limite . . . 1806
39.5 Applications . . . 1806
39.5.1 Le GPS . . . 1806
39.5.2 Les ondes électromagnétiques . . . 1806
39.6 Mécanique relativiste . . . 1806
39.6.1 Des problèmes, toujours des problèmes . . . 1807
39.6.2 Loi d'addition des vitesses . . . 1807
39.6.3 L'action d'une force . . . 1808
39.6.4 Équivalence entre la masse et l'énergie . . . 1809
39.7 Principe de correspondance . . . 1809

40 Exemples avec Sage (Vol 4) 1811

41 Épilogue : la constante de Weiner (Vol 4) 1831

42 GNU Free Documentation License (Vol 4) 1833

Bibliographie 1841

Index

λ-système, 695
p-Sylow, 319
p-groupe, 319
Éther, 1794
Événement, 1799
élément
    inversible
        dans un anneau, 142
élémentaire
    polynôme symétrique, 200
équation
    aux variations, 1467
    de Riccati, 1488
    des classes, 132
    des orbites, 132
    différentielle
        étude qualitative, 1486
        Hill, 1484
        homogène, 1487
        système, 1486
    diophantienne, 120, 158, 161
    Fredholm, 1059
    orbite-stabilisateur, 131
équation différentielle
    linéaire du premier ordre, 1498
    linéaire du premier ordre, homogène, 1499
    linéaire du second ordre, 1501
    linéaire du second ordre, homogène, 1501
    premier ordre, 1493
    second ordre, 1494
    variables séparables, 1496
équation homogène associée, 1499
équilibre
    point point une équation différentielle, 1475
équivalence
    de norme, 451
étagée
    fonction, 726
état
    apériodique, 1765
    récurrent, 1754
    récurrent positif, 1754
    transitoire, 1754
étranger
    dans leur ensemble, 166

abélianisé, 110
Abel
    angulaire, 838
    convergence radiale, 906
abscisse
    curviligne, 1143
absolument continue, 830
absorbant, 1753
accélération d'un chemin, 1127
accroissement, 652
accumulation
    dans $\mathbb{R}$, 293
    dans espace vectoriel normé, 379
action, 129
    adjointe, 130
    de groupe
        Wedderburn, 209
    domaine fondamental, 132
    fidèle, 129
    libre, 133
    transitive, 133
action de groupe, 350
    sur des matrices, 1115
adhérence, 291, 375
affine
    application, 494
    espace, 491
    sous-espace, 496
affine (application), 309
aire, 888
algébrique
    extension, 184
    nombre, 191
    par rapport à une extension de corps, 196
algébriquement
    indépendant, 242
algèbre, 148
    de parties, 690
    engendrée, 148
    polynômes, 165
algorithme, 1555
    consistant, 1555

convergeant, 1555
facteurs invariants, 163
fortement consistant, 1555
stable, 1555
alignement
dans un espace projectif, 1184
alterné
groupe, 334
polynôme, 200
alternée
forme linéaire, 343
analytique
au sens complexe, 1326
angle
d'une courbe, 1160
entre deux droites, 976
entre vecteurs, 967
orienté de vecteurs, 997
Anneau
$\mathbb{Z}/n\mathbb{Z}$
polynôme cyclotomique, 205
anneau, 93
$\mathbb{Z}/n\mathbb{Z}$, 209, 213, 329, 464
à division, 171
de séries formelles, 922
utilisation, 1344
euclidien
facteurs invariants, 163
factoriel, 153
intègre, 148
noetherien, 156
principal, 155, 175, 387
utilisation, 177
quotient par un idéal, 144
apériodique
état d'une chaîne de Markov, 1765
chaîne de Markov, 1765
application
définie positive, 357
de classe C^k, 629
différentiable, 591, 593, 626, 629, 1074, 1115
extrema lié, 1089
en escalier, 880
linéaire
théorème de Banach-Steinhaus, 1276
mesurable, 694
multilinéaire, 484
ouverte, 482
semi-définie positive, 357
tangente, 597
application réciproque, 294
approximation
de fonctions
par des polynômes, 1703
de l'unité, 1295
par polynômes, 1051
polynômiale, 1331
arc
géométriques, 1142
paramétré, 1127
arc cosinus, 964
arc sinus, 962
arc tangente, 961
archimédien, 94
associée
subdivision, 880
associés
éléments d'un anneau, 150
asymptotiquement pivotale, 1726
attractif
point fixe, 1056
automorphisme
d'espace vectoriel, 305
axiome
du choix, 89

Bézout
anneau principal, 155
calcul effectif, 117
nombres entiers, 114
polynômes, 183
Baire
espace, 298
théorème, 290, 299
tribu, 688
Banach
espace, 1249
barycentre
cas affine, 497
cas vectoriel, 985
enveloppe convexe, 501
base, 301
canonique de $\mathbb{R}^m$, 304
d'un module, 147
de Newton, 1564
de topologie, 247
dénombrable, 259
espace métrique, 259
duale, 586
espace préhilbertien, 1258
hilbertienne, 1258
utilisation, 1379
locale, 1238
Bergman (espace), 1363
Bernoulli, 1671
somme, 1689
Berry-Esséen (borne), 1670

Bessel
inégalité, 1257
biais
d'estimateur, 1716
bien
conditionné, 1536
enchaîné, 288
bijection, 294
bilinéaire, 484
binormale, 1151
birégulier
point sur une courbe, 1140
birapport, 1196
birapport dans $\mathbb{C} \cup \{\infty\}$, 1214
Bolzano-Weierstrass
espaces métriques, 283
bon
ordre, 90
borélienne
fonction, 710
tribu, 687
boréliens, 687
bord, 292
borné, 248
partie de V, 371
temps d'arrêt, 1774
bornée, 272
différentielle, 625
partie de $\mathbb{R}^m$, 535
suite, 266
boule
avec semi-normes, 295
fermée, 291, 371
ouverte, 247, 291, 371
Bruhat (décomposition), 678
Burnisde
formule, 133

Cône de lumière, 1799
canonique
base, 304
décomposition, 91
espace affine, 491
Cantor
ensemble, 762
caractéristique
d'un anneau, 145
d'une équation différentielle, 1507
polynôme, 389
sous-groupe, 107
caractère, 1029
abélien, 1023
de S_4, 1037
groupe diédral, 1041
irréductible, 1030
cardioïde, 1140
carré
dans un corps fini, 213
catégorie
ensemble de première, 290
Cauchy
critère
uniforme, 647
déterminant, 352
formule, 1326
produit, 904
théorème, 319
Cauchy-Riemann, 1321
Cauchy-Schwarz, 357, 1250
Cayley
théorème, 319
cellule d'un pavage, 879
centrale (application), 1030
centralisateur, 107, 141
centre
d'un anneau, 141
d'un groupe, 107
d'une rotation, 972
cercle
circonscrit à une courbe, 79, 1167
dans la sphère de Riemann, 1206
cercle-droite, 1207
cercles
perpendiculaires, 1177
Cesaro
moyenn, 476
chaîne, 288
de Markov, 1745
apériodique, 1765
convergence, 1766
finie, 1746
homogène, 1745
irréductible, 1751
récurrente positive, 1760
régulière, 1747
champ
conservatif, 850
de vecteurs, 801
champ dérivant d'un potentiel, 850
changement de variable, 1142
Chasles, 491
Chemin
classe C^2, 799
chemin, 799, 1127
dans $\mathbb{R}^p$, 1127
circulation, 848
clôture algébrique, 191

classe
conjugaison
dans S_4, 333
d'association, 150
de conjugaison, 109
classe C^1, 612
classe d'association, 150
codimension, 305
coefficient
de Fourier, 1367
coefficients
de Fourier, 1300
coefficients binomiaux, 143
coercion, 1270
coercive, 1104
colinéarité, 1183
combinaison
convexe, 497
combinatoire, 994
commutant, 431
commutateur
dans un groupe, 109
compacité, 283, 285, 289, 1335, 1360
sous-groupes du groupe linéaire, 680
théorème de Dini, 649
utilisation, 1109
théorème de Montel, 1361
compact, 246, 373, 537
arc paramétré, 1127
Bolzano-Weierstrass dans $\mathbb{R}^n$, 276
boule unité, 276
et fonction continue, 276, 284
fermé et borné, 275
implique fermé, 262
intervalle $[a, b]$, 274
le coup du, 829
localement, 247
opérateur, 1275
produit dénombrable, 290
produit fini, 289
quasi, 247
relatif, 1275
relativement, 247
séquentiellement, 247
suite exhaustive, 287
complément
à deux, 1526
complémentaire, 90
complété
d'un espace métrique, 1123
complétion
projective, 1186
complétude, 1121, 1123, 1286
de $\mathbb{R}$, 467
espaces L^p, 1287
complète
famille de projecteurs, 147
complet
$\mathbb{R}$, 257
corps, 94
espace mesuré, 700
espace topologique, 252
métrique, 252
composante, 652
composition
suite de, 125
concave, 1090
log-concave, 1097
condition initiale, 1494
conditionnement
absolu, 1536
d'une matrice inversible, 1570
relatif asymptotique, 1555
conjugués
éléments d'une extension, 185
connectés
indices d'une matrice, 1597, 1598
connexe
par arc, 278
connexité, 289
définition, 264
et intervalles, 273
fonction holomorphe, 1330
indice d'une courbe, 1325
le groupe $\mathrm{GL}^+(n, \mathbb{R})$, 669
par arc
fonction différentiable, 627
points d'accumulation, 285
prolongement analytique, 1125
signature d'une forme quadratique, 1112
théorème de Runge, 1331
théorème des valeurs intermédiaires, 517
utilisation
Brouwer, 1329
Conservative, 803
consistance
estimateur, 1715
ordre, 1620
constructible
angle, 237
point, 233
réel, 233
construction
des réels, 101
contenu, 167
continue, 514

fonction
en un point, 261
fonction entre espaces métriques, 279
fonction entre espaces topologiques, 261
fonction réelle, 514
forme différentielle, 586
sur espace métrique, 519
uniformément, 535
continuité, 541
fonction définie par une intégrale, 827
séquentielle, 280
sur un intervalle, 514
contraction, 1056
convergeant
schéma discret, 1620
convergence
absolue, 469
commutative, 409
dans un espace vectoriel normé, 467
de martingales, 1775
de suite, 246
en loi, 1660
en probabilité, 1660
normale, 469
ordre, 1551
presque sûrement, 1660
quadratique, 1541
rapidité, 920, 1386, 1387, 1550, 1689
série, 469
suite
dans un corps, 94
suite dans $\mathbb{R}^m$, 293
suite numérique, 266, 1051, 1371
Abel angulaire, 838
uniforme, 646
intégrale, 825
série de fonctions, 470
théorème de Dini, 649
convergent
estimateur, 1715
convexe
courbe plane, 1169
fonction sur $\mathbb{R}^n$, 1098
convexité
barycentre, 500
enveloppe de $O(n)$, 676
fonction, 1090
inégalité de Jensen, 1102
locale, 1311
méthode de Newton, 1546
stricte, 1090
utilisation, 1109
convolution, 1635, 1768
coordonnées
barycentriques, 505
cartésiennes
dans un espace affine, 504
curvilignes, 1238
cylindrique, 1020
dans un espace affine, 492
homogène, 1202
sphériques, 1020
corps, 93
archimédien, 94
complet, 94
de décomposition, 191
de rupture, 189
polynôme cyclotomique, 205
des fractions, 152
des fractions rationnelles
utilisation, 1344
extension, 193, 201
fini, 213, 217, 220
Wedderburn, 209
formellement réel, 171
ordonné, 94
premier, 173
cosinus, 951
angle entre deux vecteurs, 967
hyperbolique, 924
courbe, 1127
étude métrique, 1379
de Jordan, 1176, 1379
efficacité, 1736
fermée, 1157
simple, 1157
courbe de niveau, 571, 574
courbure, 1151
signée, 1158
totale, 1159
covariance, 1638
critère
Abel, 902
Abel pour intégrales, 826
Cauchy
uniforme, 647
de Cauchy, 467
série alternée, 474
Weierstrass, 826
série de fonctions, 898
critique
Galton-Watson, 1768
point, 1085
point d'un arc, 1140
région, 1735
valeur, 1736

cyclique
endomorphisme, 385
groupe, 107
matrice, 385
cycloïde
coordonnées normales, 1145
longueur, 1139

décalage, 1527
décimale
décomposition, 479
décomposition
Bruhat, 678
canonique, 91
corps, 191
Dunford, 416
application, 418
exponentielle de matrice, 417
Jordan
et exponentielle de matrice, 418
polaire, 674
primaire, 415
sous-espaces caractéristiques, 415
spectrale, 415
dénombrable, 92
à l'infini, 247
dénombrement, 994
partitions de $\{1, \dots, n\}$, 922
dérivé
groupe, 109
dérivée
au sens de distributions, 1427
dans Sobolev $H^1(I)$, 1427
directionnelle, 578, 582
distributionnelle, 1406
faible, 1401
fonction à valeurs dans E', 298
partielle, 578, 580, 651
seconde, 563
dérivabilité
fonction définie par une intégrale, 829
lemme de Borel, 942
dérivable, 557, 563
au sens complexe, 893
fonction, 1445
dérivation
au sens des distribution
Sobolev, 1431
déterminant, 343
Cauchy, 1054
d'un endomorphisme, 348
d'une famille de vecteurs, 345
de Cauchy, 352
et inversibilité, 348
forme linéaire alternée, 343
Gram, 352, 1054
interprétation géométrique, 776
résultant, 354, 842
utilisation, 1109
Vandermonde, 349
détermination
logarithme, 1357
principale, 1357
développable
en série entière, 929
développement
asymptotique, 646
limité
en zéro, 638
fonction holomorphe, 1321
premier ordre, 562
Taylor, 1115
degré
application $S^1 \to S^1$, 1161
d'un polynôme, 165
d'une représentation, 1023
extension de corps, 185
dense
nulle part, 290
densité, 1123
conjointe, 1634
d'une variables aléatoire, 1630
dans un espace de fonction
critère de Weyl, 1371
de $\mathbb{Q}$ dans $\mathbb{R}$, 258
utilisation, 1096
de $\mathrm{GL}(n, \mathbb{R})$ dans $\mathbb{M}(n, \mathbb{R})$, 669
de $\mathscr{D}(\mathbb{R}^n)$ dans $L^1(\mathbb{R}^n)$, 1318
de $C_c^\infty(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$, 1291
de $L^2\big([0,1]\big)$ dans $L^p\big([0,1]\big)$, 1292
de $S^+(n, \mathbb{R})$ dans $S^{++}(n, \mathbb{R})$, 673
des fonctions étagées dans L^p, 1291
des polynômes
dans $C_c^0[0,1]$, 1703
matrices diagonalisables dans $\mathbb{M}(n, \mathbb{C})$, 669
mesure, 78, 739
points extrémaux dans $\mathcal{L}$, 675
prolongement, 1121
densité d'une mesure, 739
diédral, 995
diagonale
dominante, 1602
fortement dominante, 1602
strictement dominante, 1602
diagonalisable, 397
et polynôme minimum scindé, 397
exponentielle, 418

diagonalisation
cas complexe, 404
cas réel, 406
endomorphisme auto-adjoint, 446
simultanée, 399
diamètre, 535
difféomorphisme, 616, 862, 1017
de classe C^k, 629
différence
centrée, 1609
divisée, 1565
progressive, 1609
régressive, 1609
différentiabilité, 627
différentiable, 597
dans un Banach, 612
deux fois, 629
sur un ouvert, 603
différentielle, 591, 592, 612
de $u \mapsto u^{-1}$, 621
partielle, 617
totale, 652
dilatation, 662
dilatation (matrice), 162
dimension, 304
n-formes multilinéaires alternées, 343
définition, 304
sous espace affine, 496
utilisation, 501
direction, 579, 1148
sous-espace affine, 496
Dirichlet
noyau, 1368
théorème, 1368
théorème (sur les nombres premiers), 209
disque de convergence, 901, 903
distance, 247
associée à une norme, 251
compatible, 249
entre deux mesures de probabilités, 1731
invariante, 249
point et ensemble, 282
distingué, 107
distribution, 1404
équation de Schrödinger, 1490
de Dirac, 1410
produit par une fonction, 1406
tempérée, 1410
divergence, 655
diviseur
dans un anneau, 141
de zéro, 148
de zéro à droite, 148
polynôme, 166
division
euclidienne, 113, 166
harmonique, 1217
domaine, 511
fondamental d'une action, 132
dominé
modèle statistique, 1714
dominée
convergence (Lebesgue), 738
mesure, 743
droite
dans la sphère de Riemann, 1206
projective, 1183
dual
algébrique, 312
d'un espace de Hilbert, 1255
de $\mathbb{M}(n, \mathbb{K})$, 316
de $L^p(\Omega)$, 1309
de L^p, 1305
topologique, 312
Dunford
décomposition, 416
dyadique, 765

écart-type, 1637
échantillon, 1708, 1710
effectif
empirique, 1739
efficacité
courbe, 1736
d'une méthode itérative, 1559
élément
de surface, 811
de torsion, 136
élément de surface, 793
ellipsoïde, 445
elliptique
équation aux dérivées partielles, 1516
endomorphisme, 305
auto-adjoint, 446
cyclique, 385
décomposition
polaire, 674
diagonalisable, 437, 671, 1486
Dunford, 416
diagonalisation, 406
nilpotent
Dunford, 416
préservant une forme quadratique, 682
sous-espace stable, 416, 1486
engendré, 144
λ-système, 695
corps et anneau, 186

idéal dans un anneau, 144
sous-espace affine, 496
sous-groupe, 108
tribu, 686
par une variable aléatoire, 1631
ensemble
de Cantor, 762
différence symétrique, 91
infini, 89
entrelacement, 1030
enveloppe
convexe, 500
équation
différentielle
linéaire, 1448
ordinaire d'ordre 1, 1445
variables séparées, 1450
générale de degré n, 243
équicontinuité, 1275
équi-intégrable, 1775
équivalence
arcs paramétrés, 1142
chemin, 1134
classe de fonctions, 1278
de représentations, 1030
de suites, 468
homotopie, 1325
norme, 451
relation, 91
suite de composition, 127
erreur, 1566, 1596
assignation, 1530
de consistance, 1620
discrétisation, 1620
quadratique, 1569
troncature, 1530
erreur relative, 1530
escalier, 726
espérance, 1636
conditionnelle, 1640, 1641
événement, 1642
variable aléatoire, 1646
Espace
de Sobolev, 1433
espace
L^2
Sobolev, 1431
L^p, 1279
Banach, 1249
de Baire, 298
de Bergman, 1363
de fonctions
L^p, 1287
Sobolev H^1, 1431
de Hilbert
espace de Sobolev H^1, 1431
de probabilité, 1629
de Schwartz, 1315, 1410
de Sobolev, 1427, 1434
euclidien, 361
métrique, 247
base de topologie, 259
mesuré, 692
complété, 703
mesurable, 685
projectif, 1183
propre, 380
tangent, 1118
topologique, 264
métrisable, 279
vectoriel
dimension, 343
espace vectoriel
topologique, 249
estimateur, 1715
biais, 1716
consistant, 1715
convergent, 1715
de fonction de répartition, 1721
maximum de vraisemblance, 1718
estimation
des grands écarts, 1689
étrangers
polynômes, 166
Euclide
algorithme étendu, 116
lemme, 118
euclidien
anneau, 156
espace, 358
Euler
indicatrice, 139
événement, 1629
exact
intervalle de confiance, 1724
excès
intervalle de confiance, 1724
exhaustive (suite de compacts), 287
exponentielle, 916
complexe, 1347
convergence, 476
de matrice, 418, 486, 938
utilisation, 1082
existence, 913
rapide, 466
unicité, 913

exposant, 326, 437
d'un groupe, 108
extension
corps de base, 420
de corps, 184, 201
algébrique, 184, 189
finie, 222
simple, 186
utilisation, 238
isométrie, 1122
extrémal
point dans un convexe, 675
extrémité
d'un intervalle, 273
extrapolation, 1564
extrema, 1086
lié, 1089
local
relatif, 1088
volume d'un ellipsoïde, 1109
extremum, 1115

facteur
intégrant, 1489
factoriel
anneau, 153
factorisation
de polynôme, 168, 187
faisceau de droites, 1203
famille
sommable, 410
Fatou, 737
Fejér
noyau, 1368
fermé, 245, 292, 373
dans un compact, 262
fermeture, 375
séquentielle, 281
fidèle (action), 129
filtration, 1771
fine
subdivision, 1128
fixateur, 129
flot, 1462, 1507
flux
d'un champ de vecteur, 818
flux d'un champ de vecteurs, 815
fonction, 511
Γ d'Euler, 1341
étagée, 1290
à décroissance rapide, 1315
borélienne, 710
caractéristique, 882
d'une variable aléatoire, 1656
continue
égales, 282
par morceaux, 729
convexe, 1090, 1096
croissante, 511
décroissante, 511
définie par une intégrale, 827, 832, 942, 1360
Γ d'Euler, 1341
utilisation, 1676
de classe $\mathcal{C}^1$, 625
de Dirichlet, 1373
de Möbius, 231
de répartition, 1656
en escalier intégrable, 881
génératrice, 1657
holomorphe, 893, 1360
théorème de Montel, 1361
image, 511
méromorphe
Γ d'Euler, 1341
monotone, 512
par morceaux, 729
valeurs vectorielles, 652
fonctionnelle
énergie, 1272
fondamental
domaine d'une action, 132
forme
bilinéaire, 357
non dégénérée, 443
différentielle, 586
exacte, 835
fermée, 835
linéaire
différentielle, 1089
quadratique, 441, 988, 1112, 1114, 1115
groupe orthogonal, 682
matrice associée, 441
forme canonique
fonction simple, 726
forme linéaire
alternée, 343
formule
Bayes, 1639
Burnside, 133
d'expulsion (produit vectoriel), 368
de Cauchy, 1326
Hadamard, 903
inversion Möbius, 231
probabilité totales, 1639
sommatoire de Poisson, 1386
Stirling, 468
Taylor

reste intégral, 637
utilisation, 1550
Fourier, 1386
série
utilisation, 1379
transformée
groupe abélien fini, 1025
fréquence
empirique, 1739, 1757
fraction
rationnelle
intégration, 842
fraction rationnelle, 872
fractions (corps), 152
Fredholm
équation, 1059
Frenet
formules, 1152
Fresnel
intégrale, 1349
Frobénius
réduction, 428
Frobenius
morphisme, 146
frontière, 292, 378
Fubini
théorème
dans $\mathbb{R}^n$, 791

générateur, 107
génératrice
partie d'un module, 147
géométrie
avec des groupes, 988, 1224
avec nombres complexes, 988, 1224
géométrique
avec des nombres complexes, 1379
Galton-Watson
sous-critique, 1768
sur-critique, 1768
Gauss
lemme
polynômes, 184
somme de, 215
Gershgorin
disque, 1599
Grönwall (lemme), 1445, 1446
gradient, 594, 603, 624
Gram (déterminant), 352
Gram-Schmidt, 361
graphe, 511, 560
de transition (chaîne de Markov), 1751
fonction, 570
fonction de deux variables, 573
groupe
p-groupe, 319
$\mathrm{GL}(n,\mathbb{R})$, 1114
action, 1224
utilisation, 682
agissant sur un ensemble
diédral, 988
alterné, 334
circulaire, 1221
dérivé, 109
de $\mathrm{GL}(n,\mathbb{K})$, 667
de $\mathrm{SL}(n,\mathbb{K})$, 667
du groupe alterné, 337
du groupe symétrique, 335
de Galois, 242
de permutation, 1041
caractères de S_4, 1037
de permutations, 994
de torsion, 136
des isométries
espace métrique, 248
des symétries, 986
diédral, 988, 995
générateurs (preuve), 988
générateurs (utilisation), 1041
en géométrie, 988
et géométrie, 343, 994, 1224
isométries du cube, 462
fini, 209, 213, 321, 329, 464, 994
alterné, 336
diédral, 988
Wedderburn, 209
linéaire, 678
décomposition polaire, 674
enveloppe convexe de $\Omega(n)$, 676
hyperplan, 317
sous-groupes compacts, 680
modulaire, 1224
orthogonal, 363
d'une forme quadratique, 682
partie génératrice, 336, 464, 1224
permutation, 343, 350, 464, 678
diédral, 988
projectif, 1190
quotient, 125
résoluble, 128
spécial orthogonal, 364
symétrique, 134
action sur un triangle, 1027
groupe dérivé
de $\mathrm{GL}(n,\mathbb{C})$, 661

Hadamard
formule, 903

Hardy-Littlewood (théorème), 1051
Hausdorff, 246
Heine (théorème), 536
hermitien
produit scalaire, 362
hessienne, 631
Hilbert, 1249
holomorphe, 893
homéomorphisme, 260
homogène
chaîne de Markov, 1745
homographie, 1189, 1224
sur $\mathbb{C} \cup \{\infty\}$, 1210
homotopie, 1163
hyperbolique
équation aux dérivées partielles, 1516
hyperplan, 981
de $\mathbb{M}(n, \mathbb{K})$, 317
sépare
au sens strict, 1311
séparer
au sens large, 1311
hypothèse
alternative, 1735
composite, 1735
multiple, 1735
nulle, 1735
simple, 1735

idéal
bilatère, 94
dans un anneau, 93
maximal, 153
maximum, 197
principal
à droite, 153
à gauche, 153
identifiable, 1714
identité
polarisation, 441
image, 511
inégalité
arithmético-géométrique, 1102
Bessel, 1257
Cauchy-Schwarz, 357, 1250
de Khintchine, 1686
de la moyenne, 626
des pentes, 1091
Hölder, 1283
utilisation, 1637, 1703
isopérimétrique, 1379
Jensen, 1102
espérance conditionnelle, 1655
pour une somme, 1102
version intégrale, 1282
Kantorovitch, 1103
Markov, 1665
Minkowski, 1284
triangulaire, 247, 250
produit scalaire, 358
incompressible
champ de vecteur, 657
indécomposable
module, 147
indépendance, 1633
événements, 1630
utilisation, 1701, 1778
affine, 504
algébrique, 242
projective, 1191
sous tribus, 1630
variables aléatoires, 1631
indicatrice d'Euler, 139
indice, 123
d'une courbe dans $\mathbb{C}$, 1325
de rotation, 1162
inductif, 90
induite
topologie, 375
tribu, 686
inférence statistique, 1707
infimum, 269
injection, 294
intégrable, 730
fonction non en escalier, 887
fonction positive, 892
intégrale
calcul, 1371
convergente, 774, 842
d'une fonction sur une carte, 794
d'une fonction sur une variété, 799
d'une forme différentielle, 802
fonction en escalier, 881
Fresnel, 1349
impropre, 773, 774
sur un chemin, 799
intégration
fraction rationnelle, 842
intérieur, 291
d'un ensemble, 372
point, 372
intègre
anneau, 148
interpolation, 1564
Intervalle, 1799
intervalle, 90, 511
fermé, 511

longueur, 753
ouvert, 511
intervalle de confiance
asymptotique, 1728
invariant
de similitude, 428
invariante
mesure
pour une chaîne de Markov, 1762
inverse généralisé, 1693
inversion, 1178
dans le groupe symétrique, 332
inversion dans $\mathbb{C} \cup \{\infty\}$, 1208
involution, 399
irréductible
chaîne de Markov, 1751
dans un anneau, 152
module, 147
polynôme, 180
représentation, 1028
irrationalité
$\sqrt{2}$, 97
isobarycentre, 498
isolé
élément de $\mathbb{R}$, 293
point dans un espace vectoriel normé, 379
isométrie
d'espaces métriques, 248
de forme quadratique, 945
de l'espace euclidien $\mathbb{R}^2$, 988
espace euclidien
isométries du cube, 462
forme bilinéaire, 358
groupe, 248
isomorphisme
$(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}$, 213
d'anneaux, 93
d'espaces topologiques, 260
de corps, 93
espace affine, 495
espace vectoriel normé, 459
espaces vectoriels, 305
isotrope
cône, 444
totalement, 444
isotrope (vecteur), 444

jacobien, 624, 654
Jordan
chemin, 818
courbe, 1379
réduction, 430
Jordan-Hölder, 125

Kronecker, 304

lacet, 1325
Lagrange
multiplicateur, 1089
polynôme, 316
théorème, 123
lagrangien, 1089
Laplace
transformée, 1657
laplacien, 1244
Legendre
symbole, 213
Leibnitz, 559, 562
applications entre espaces vectoriels normés, 617
lemme
Borel, 942
d'Euclide, 118
de Borel-Cantelli, 1663
de Gauss
contenu de polynôme, 221
pour des entiers, 118
de Morse, 1115
de Schreider, 127
de Slutsky, 1662
de transfert, 711
de Zorn, 90
des noyaux, 381
Fatou, 737
Gauss
dans un anneau principal, 155
polynômes, 184
Grönwall, 1445, 1446
Hadamard, 834
regroupement, 1633
Schur complexe, 401
Schur réel, 405
Levi-Civita, 1239
libre
action, 133
partie, 301
partie d'un module, 147
Ligne d'univers, 1795
limite, 413
d'ensembles, 694
d'une fonction, 260
de fonctions holomorphes, 1360
de suite
espace topologique, 246
fonction, 260, 539
fonction de plusieurs variables, 540
inférieure, 719, 720
inversion, 922, 1051, 1279, 1360

permutation
utilisation, 1658
suite, 467
suite dans $\mathbb{R}^m$, 293
suite numérique, 266
supérieure, 719, 720
unicité, 260
linéaire
application, 305
Lipschitz, 1056
localement, 628
Lipschitzienne, 627
localement
intégrable, 773
log-concave, 1097
logarithme
complexe, 1354
dans $\mathbb{C}$, 1351
de matrice, 939
sur les réels positifs, 917
loi
χ^2, 1687
binomiale
comportement asymptotique, 1689
conjointe, 1633
d'une variable aléatoire, 1658
de Poisson, 1674
des grands nombres
forte, 1666
pour les chaînes de Markov, 1764
processus de Poisson, 1786
utilisation, 1689, 1701
marginale, 1633
normale
vecteur gaussien, 1681
parente, 1707, 1708
parente d'un échantillon, 1710
réciprocité quadratique, 217
sans mémoire, 1676
Student, 1687
Lokta-Voltera, 1479
longueur
élément de, 1138
arc géométrique, 1143
d'un arc paramétré compact, 1128
d'un intervalle, 753
d'une arrête, 878
longueur d'arc, 1130

M-matrice, 1604
méromorphe, 1340
Méthode
de Newton, 1545
méthode
des chemins, 548
Newton, 1550
cas convexe, 1546
métrique
discrète, 248
métrisable, 254
maigre, 687
maigre (ensemble), 290
majorant, 268, 269
essentiel, 1281
Markov
inégalité, 1665, 1713
martingale, 1771
bornée dans $L^2(\Omega)$, 1772
matrice, 678, 1224
équivalence, 307
dans le groupe linéaire, 682
associée à une forme quadratique, 441
compagnon, 426
creuse, 1574
cyclique, 385
de dilatation, 162
de permutation, 162
de similitude, 893
de Sylvester, 352
de transition, 1745
de transvection, 162
dense, 1574
hermitienne
racine carré, 671
jacobienne, 603, 624
normale, 404
orthogonale, 363
permutation
élémentaire, 1581
réductible, 1598
racine carré, 671
semblable, 671
semblables, 988, 1114
stochastique, 1745
symétrique, 1114
réelle, 1112
trigonalisable, 400
matrices
similitude, 308
maximal
idéal, 153
maximale
partie orthonormale, 1259
maximum, 268, 271
global, 1084
local, 1084
mesurable

application, 694
au sens de m^*, 699
ensemble, 692
fonction, 710
Lebesgue, 877
mesure
σ-finie, 692
absolument continue, 743
angle entre vecteurs, 998
complexe, 744
dans une carte, 793
de Borel, 716
de comptage, 693, 790
de Haar, 1062, 1063
de Lebesgue, 756
de Radon, 716, 1338
extérieure, 689
externe, 877
finie, 692
sur algèbre de partie, 690
image, 713
positive, 692
probabilité, 1629
produit, 750
régulière, 716
extérieure, 716
intérieure, 716
sur algèbre de partie, 690
minimum, 271
ensemble ordonné, 89
minorant, 269
modèle
échantillonnage, 1708, 1710
paramétrique, 1708
statistique, 1707
modulaire (groupe), 1224
module
de continuité, 1120
indécomposable, 147
irréductible, 147
simple, 147
sur un anneau, 146
moment, 1636
fonction génératrice, 1657
monôme, 165
monogène, 107
extension de corps, 186
monotonie, 1483
morceau
fonction continue ou monotone, 729
morphisme
d'algèbres, 148
d'anneaux, 93
de corps, 93
espace vectoriel normé, 459
Frobenius, 146
moyenne
de Cesaro, 476
empirique, 1638
empirique d'un échantillon, 1711
quadratique, 1637
multiplicateur
de Lagrange, 1089
multiplicité
racine d'un polynôme, 168
racine de $f(x) = 0$, 1544
valeur propre
algébrique, 390
géométrique, 390

négatif, 103
négligeable
partie d'un espace mesuré, 700
nabla, 611
Newton
méthode, 1550
nilpotent, 142
niveau de confiance, 1724
nombre
complexe
norme 1, 209
dénormalisé, 1528
de Fermat, 238
de Néper, 914
normal, 1701
normalisée, 1528
premier, 113, 175, 209, 213, 329, 464, 920
dans leur ensemble, 113
deux nombres entre eux, 113
théorème des deux carrés, 177
tours d'une courbe plane, 1161
nombre premier
polynôme cyclotomique, 205
non dénombrable, 92
normé
espace vectoriel, 250
normal
arc paramétré, 1143
endomorphisme, 404
nombre, 1701
sous-groupe, 107
normal extérieur
vecteur, 817
normale
loi réduite, 1680
principale, 1151
normalisateur, 107

norme, 250
  équivalence, 451
  d'algèbre, 455
  d'application linéaire, 453
  d'une application linéaire, 453
  euclidienne, 251
    dans $\mathbb{R}^m$, 449
  supremum, 251
  vecteur, 359
noyau
  d'une forme bilinéaire, 444
  Dirichlet, 1368
  Fejér, 1368
nulle part dense, 290

observation, 1629
opérateur
  adjoint, 362
  autoadjoint, 363
  définit positif, 406
  hermitien, 363
opérateurs
  compatibles, 1269
opposés
  chemins, 1134
orbite
  d'un point sous une action, 129
ordre, 89
  élément, 108
  bon ordre, 90
  d'un groupe, 108
  d'un polynôme, 167
  dans un corps, 94
  de convergence d'un schéma, 1620
  distribution, 1407
  partiel, 89
  sur un anneau factoriel, 153
  total, 89
orientable
  variété, 796
orientation, 795
origine
  abscisse curviligne, 1143
  repère affine, 504
orthogonal, 448, 1254
  coordonnées curviligne, 1238
  famille de projecteurs, 147
  matrice, 363
  opérateur, 363
  sous-espace, 312
  vecteur, 359
orthonormé, 449
  système, 1256
oscillation
  d'une fonction, 530
  d'une fonction en un point, 530
osculateur (cercle), 1156
ouvert, 245, 272, 292, 373
  dans $\mathbb{R}^n$, 291

pôle, 1005
parabolique
  équation aux dérivées partielles, 1516
parallèle
  sous-espaces affines, 496
paramétrages
  admissible, 1142
paramétrisation, 1134
  normale, 1143
Parseval, 1263
partie
  génératrice, 301
  régulière, 638
  totale, 1256
partie génératrice, 107
partition
  d'un entier en parts fixées, 1344
  dénombrable mesurable, 722
  de l'unité, 798
pavé, 877, 878
pavable, 878
Pearson
  theoreme, 1740
peigne de Dirac, 1416
permutation, 134
  matrice, 162
permuter
  dérivée et intégrale
    $\mathbb{R}^n$, 832
    dans $\mathbb{R}$, 829
    dans $\mathbb{R}$ avec les bornes, 831
  dérivée et limite, 649
  différentielle et intégrale
    $\mathbb{R}^n$, 833
  intégrale
    et série, 790
  limite et intégrale, 827
    convergence dominée, 738
    convergence monotone, 735
    espace mesuré, 827, 828
  série entière et dérivation, 908
  série entière et intégration, 910
  somme et intégrale, 736, 899
petit théorème de Fermat, 173
PGCD
  dans un anneau intègre, 150
  polynômes, 183
pgcd, 113, 142

calcul effectif, 116
pivotale, 1726
plan
projectif, 1183
tangent, 612, 652
Plancherel, 1263
plongement, 1121
Poincaré (demi-plan), 1224
point
d'équilibre
stable, 1475
pondéré, 497
point critique
définition, 1116
point fixe, 1768
attractif, 1056
Brouwer, 1059
Picard, 1057
Schauder, 1061
Poisson
formule sommatoire, 1386
processus, 1785
polarisation (identité), 441
polynôme
à plusieurs indéterminées, 220, 355
alterné, 200
annulateur, 381, 384
caractéristique, 389, 403
contenu, 167
cyclotomique, 203
irréductibilité, 205
propriétés, 205
d'endomorphisme, 671
décomposition de Dunford, 416
de Bernstein, 1703
irréductible, 180
séparable, 192
sur $\mathbb{F}_q$, 232
Lagrange, 316
minimal, 188, 383
d'un élément d'une extension, 184
ponctuel, 384
racines, 201
séparable, 192
scindé, 181
semi-symétrique, 200
symétrique, 200, 201, 220, 350, 355
élémentaire, 200, 202
trigonométrique, 1298
portée
mesure, 743
positif, 103
potentiel, 806, 850
PPCM
dans un anneau intègre, 150
ppcm, 113
précision
simple, 1527, 1528
préhilbertien, 1249
premier
corps, 173
deux éléments d'un anneau principal, 154
deux polynômes entre eux, 167
idéal, 154
sous corps, 173
premier temps d'atteinte, 1753
premier type
région solide, 890
presque
nulle, 136
partout, 694
surjective, 1312
primitif
élément d'un corps, 213
élément d'une extension de corps, 186
polynôme, 167
au sens du pgcd, 167
racine, 225
triplet pythagoricien, 159
primitive, 557, 848
de fonction continue, 1049
et intégrale, 770, 774
fonction, 569
principal
anneau, 154
idéal, 153
principe
prolongement analytique, 1125
zéros isolés, 1330
Principe de correspondance, 1810
probabilité
conditionnelle, 1639
problème
aux limites d'évolution, 1517
aux limites stationnaires, 1517
bien posé, 1517
limite de Dirichlet, 1517
limite de Von Neumann, 1517
problème de Cauchy, 1494
processus
adapté à une filtration, 1771
arrêté, 1776
croissant prévisible, 1776
Galton-Watson, 1766
Poisson, 1785
sans mémoire, 1676

produit
d'espaces vectoriels normés, 480
d'une mesure par une fonction, 78, 739
de Cauchy, 904
de convolution, 1294
et Fourier, 1383
distribution et fonction, 1406
espaces mesurés, 752
espaces topologiques, 246
mixte, 367, 369
scalaire
en général, 357
hermitien, 362
sur $\mathbb{M}(n, \mathbb{R})$, 461
sur $\mathbb{R}^n$, 447
semi-direct, 134
tensoriel
de représentations, 1036
vectoriel, 367, 368
produit remarquable, 1333
projecteur
dans un module, 147
projectif
complétion, 1186
droite, 1183, 1184
espace, 1183
groupe, 1190
hyperplan, 1184
plan, 1183
repère, 1191
sous-espace, 1183
projection
orthogonale, 1252
prolongement
analytique, 1125
utilisation, 1362
de fonctions, 1121
lemme de Borel, 942
méromorphe de la fonction Γ, 1341
par continuité, 526, 529
dans $H^1(I)$, 1430
par densité, 1121
théorème de Hahn, 708
propriété d'intersection non vide, 261
puissance
d'un point, 1177
d'un test, 1736
d'une inversion, 1208

quasi-compact, 247
quaternion, 196
quotient, 113, 166
dans une suite de composition, 125
de groupe, 125
de groupes, 328

réciproque
continuité, 294
dérivabilité, 561
récurrent
état, 1754
nul, 1754
point d'un système dynamique, 718
positif, 1754
réduction
d'endomorphisme, 416
Frobénius, 428
Jordan, 430
réel, 101
réflexif, 1309
réflexion, 949
dans $\mathbb{R}^2$, 969
glissée, 978
par rapport à un hyperplan, 982
région
critique, 1735
de confiance exact, 1726
de rejet, 1735
régularité
d'une mesure, 716
extérieure de la mesure de Lebesgue, 759
intérieure de la mesure de Lebesgue, 761
régulière
surface, 807
régulier
arc, 1140
chemin, 818
point d'un arc, 1140
polyèdre, 950
régulier à droite, 141
répulsif
point fixe, 1056
résidu
méthode itérative, 1596
résolvante, 1456
résultant, 353
utilisation, 355, 842
règle
du produit nul, 148
Règle de Leibnitz, 562
racine
carré
de matrice hermitienne, 671
carré de matrice
hermitienne positive, 671
d'un polynôme, 167
de l'unité, 136, 205, 209, 437, 988, 1224
primitive, 138

utilisation, 205
de polynôme, 188
multiple, 1544
primitive, 225
simple, 1544
racine carré, 518
raffinement, 1128
subdivision d'un pavé, 879
rang, 306, 343, 988
classe d'équivalence, 308
diagonalisation, 406
différentielle, 1089
utilisation, 1255
rare, 687
rayon
de convergence, 901
de courbure, 1151
de torsion, 1152
spectral, 417, 455
recouvrement, 537
rectangle
produit de tribus, 745
rectifiable, 1128
arc géométrique, 1143
rejet
région dans une prise de décision, 1735
relèvement, 1161
relations
coefficient-racines, 202
de Chasles, 491, 866
relativement
compact, 247, 1275
repère
affine, 504
cartésien
espace affine, 491
de Frenet, 1151
projectif, 1191
Représentation
virgule flottante normalisée, 1527
représentation
de groupe fini
caractères de S_4, 1037
fidèle, 1023
groupe diédral, 1041
irréductible, 1028
produit tensoriel, 1036
régulière gauche, 1032
virgule fixe, 1527
reste, 113, 166
d'un développement limité, 638
risque
première espèce, 1736
quadratique, 1715
seconde espèce, 1736
rotation
en dimension 2, 971
rotation-homothétie, 1208
rupture
corps, 189

séparé, 246
espace topologique, 246
séparable, 1256
élément d'une extension, 194
espace topologique, 246
extension de corps, 194
polynôme irréductible, 192
polynôme non constant, 192
sépare
les points, 1046
série
dans un espace vectoriel normé, 469
de Fourier, 1374, 1386
utilisation, 1379
de Laurent, 1340
de puissance, 901
divergence, 469
donnant $(1-A)^{-1}$, 471
entière, 901, 922, 1386
Abel angulaire, 838
fonctions holomorphes, 1326
processus de Markov, 1768
utilisation, 1344, 1658
fonctions, 1051, 1386
génératrice d'une suite, 929
utilisation, 1344
géométrique, 475
harmonique, 475
nombres, 1051
numérique, 920, 922
Riemann, 475
Taylor, 930
schéma
consistant, 1620
schéma numérique, 1620
Schrödinger, 1490
Schur (théorème), 1030
section, 580
de graphe, 571
propriété des, 747
segment
dans $\mathbb{R}^p$, 273
dans un espace affine, 498
semblables
matrices, 393
semi-définie positive, 406

semi-norme, 295
semi-simple
endomorphisme, 386
semi-symétrique
polynôme, 200
signature
d'une permutation, 332
similitude, 893
simple
extension de corps, 186
fonction, 726
groupe, 107
module, 147
singularité, 1339
effaçable, 1339
pôle, 1339
sinus, 951
hyperbolique, 924
sinus cardinal, 875
solfège, 137
solution
générale, 1493
particulière, 1493
somme
inférieure, 884
partielle, 469
supérieure, 884
somme directe (de représentations), 1027
somme partielles
Abel angulaire, 838
sommet, 1175
sous anneau, 144
sous arc, 1127
sous-additivité
sur algèbre de parties, 690
sous-espace
affine engendré par une partie, 496
caractéristique, 414
sous-groupe
caractéristique, 107
distingué, 329
dans le groupe alterné, 336
engendré, 108
normal, 125, 328
sous-martingale, 1771
sous-module, 147
sous-suite, 535
spectre
d'un endomorphisme, 380
matrice hermitienne, 403
matrice symétrique réelle, 406
sphère, 371
de Riemann, 1205
stabilisateur, 129
stabilité
d'un point d'équilibre, 1475
Lyapunov, 1475
stable, 1535
schéma numérique, 1623
stathme
sur $\mathbb{Z}[i]$, 174
stathme euclidien, 156
stationnaire
chaîne de Markov, 1762
statistique, 1715
statistiques
descriptives, 1707
strictement
convexe
sur $\mathbb{R}^n$, 1098
structure
complexe, 1158
structure d'anneau canonique, 141
Student, 1687, 1728
subdivision, 879
associée à une fonction, 880
d'un intervalle, 1128
subordonnée
norme, 453
suite, 136
équirépartie, 1371
critère de Weyl, 1371
arithméticogéométrique, 476
définie par itération, 1550
de Cauchy, 252
dans un corps, 94
de composition, 125
de fonctions, 1279
théorème de Montel, 1361
de fonctions intégrables, 827, 1360
de Jordan-Hölder, 125
exacte, 134
régularisante, 1393
support, 881
d'une permutation, 332
distribution, 1407
famille d'éléments, 136
supremum, 268, 269
d'une suite d'ensembles, 685
sur-martingale, 1771
surface paramétrée, 806
surjection, 294
Sylow
p-Sylow, 319
Sylvester (matrice), 352
symétrique

polynôme, 200
symbole
de Legendre, 213
symbole principal, 1507
système
fondamental, 1453
orthonormé, 1256
trigonométrique, 1257, 1297

tétraèdre, 950
tangent
vecteur unitaire, 1151
tangente, 960, 1148
tangente à un chemin, 1118
Tangente hyperbolique, 1804
tangente hyperbolique, 925
taubérien, 1049
taux d'accroissement, 1091
Taylor, 634
série entière, 929
temps d'arrêt, 1774
temps de retour, 1753
terminée
martingale, 1775
test, 1735
bilatéral, 1736
unilatéral, 1736
théorème
élément primitif, 195, 222, 223
accroissements finis, 619
dérivée directionnelle, 582
dans $\mathbb{R}$, 566
forme générale, 626
Ascoli, 1275
Bézout
polynômes, 183
utilisation, 353
Baire, 290
Banach-Steinhaus, 1276
avec semi-normes, 1277
base incomplète, 303
Beppo-Levi, 735
Bolzano-Weierstrass, 283
Bolzano-Weierstrass dans $\mathbb{R}^n$, 276
Borel-Cantelli, 1629
Borel-Lebesgue, 275
Brouwer, 1060
dimension 2, 1329
Carathéodory, 501
Cauchy
groupe, 124
Cauchy-Arzela, 1071
Cauchy-Lipschitz, 1064
Cayley-Hamilton, 392, 670
central limite, 1667
processus de Poisson, 1787
Chevalley-Warning, 220
chinois, 179
anneau des polynômes, 180
anneau principal, 154
Cochran, 1712
Cochrane, 1713
convergence
dominée de Lebesgue, 738
monotone, 735
d'Alembert-Gauss, 180
décomposition des noyaux
et exponentielle de matrice, 383
de Baire, 299
de Jordan, 1176
de représentation de Riesz, 1255
des deux carrés, 177
version faible, 175
Dini, 649
Dirichlet, 1368
forme faible, 209
Doob, 1775
du rang, 306
extension d'isométrie, 1122
extrema
lié, 1089
Fejér, 1369
fonction implicite dans $\mathbb{R}^n$, 1077
fonction implicite dans Banach, 1076
fondamental du calcul intégral, 770
Fubini
dans $\mathbb{R}^n$, 791
espace mesuré, 786
version compacte dans $\mathbb{R}^2$, 852
Fubini-Tonelli, 783
fuite des compacts, 1459
Gauss
polynômes, 183
Gauss-Wantzel, 238
Glivenko-Cantelli, 1721
Hahn-Banach, 1310
Hardy-Littlewood, 1051
Heine, 536
incidence, 1184
inversion locale, 1074
utilisation, 1089, 1115
isomorphisme
second, 111
troisième, 112
isomorphisme de Banach, 1275
Jordan, 1379
Kronecker, 355

Lagrange, 123
Lie-Kolchin, 439
Lokta-Voltera, 1479
Markov-Takutani, 1063
Montel, 1361
Pappus
affine, 1188
projectif, 1188
Pearson, 1740
petit de Fermat, 173
Picard, 1057
point fixe
Brouwer, 1329
projection
cas vectoriel, 1252
partie fermée convexe, 1250
prolongement de Hahn, 708
prolongement de Riemann, 1340
Radon-Nikodym, 743
complexe, 744
représentation de Riesz, 1309
Rolle, 565
Rothstein-Trager, 842
Runge, 1331
Schauder, 1061
Schur, 1030
spectral, 415
autoadjoint, 446
matrice symétrique, 406
matrices normales, 404
stabilité de Lyapunov, 1475
Stone-Weierstrass, 1046, 1048
Sylvester, 988
taubérien, 1049
taubérien faible, 840
transfert, 1659
Tykhonov, 288
dénombrable, 290
fini, 289
valeurs intermédiaires, 517
Von Neumann, 1082
Wedderburn, 209
Weierstrass, 284
topologie, 245, 375
$*$-faible, 297, 1405
p-adique, 488
et semi-normes, 295
faible, 455
forte, 454
induite, 263
métrique, 248, 272
produit, 246
sur $\mathscr{D}(\Omega)$, 1402
sur $\mathscr{D}(K)$, 1402
sur $C^\infty(\Omega)$, 1402
sur dual topologique, 297
usuelle sur $\mathbb{R}^n$, 375
topologique
somme directe, 1254
torsion, 1152
d'un groupe, 136
totale, 1256
trace, 1437
dual de $\mathbb{M}(n,\mathbb{K})$, 316
endomorphisme, 394
matrice, 394
produit scalaire sur $\mathbb{M}(n,\mathbb{R})$, 461
unicité pour la propriété de trace, 317
transcendant, 196
par rapport à une extension de corps, 196
transformée
de Cauchy, 1338
de Fourier, 1383, 1656
continuité, 1384
groupe abélien fini, 1025
Fourier
distribution tempérée, 1412
Laplace, 1657, 1658
transformation
Fourier, 1386
gaussienne, 1575
transient
état, 1754
transition
probabilité, 1745
transitive, 133
transitoire
état, 1754
transposée, 313
transvection, 662
transvection (matrice), 162
transversale, 132
tribu, 685
borélienne, 687
de Baire, 688
de Lebesgue, 756
engendrée, 686
par un événement, 1631
par une application, 711
par une variable aléatoire, 1631
induite, 686
produit, 745
trigonalisation
et polynôme caractéristique, 400
simultanée, 439
triplet

pythagoricien, 159
type
fini
en algèbre, 196
espace vectoriel, 302

unicité
des mesures, 697
unipotent, 142
unitaire
normale principale, 1151

valeur
principale (distribution), 1410
propre, 380
singulière, 420
valeur absolue
p-adique, 488
valeur principale, 1351
valeur propre
d'une forme quadratique, 443
valuation
p-adique, 488
d'un polynôme, 166
Vandermonde (déterminant), 349
variété, 1089
variété
orientée, 796
variable
de décision, 1736
variable aléatoire, 1630
absolument continue, 1630
Bernoulli
marche aléatoire, 1748
utilisation, 1703
binomiale
utilisation, 1778
centrée, 1636
de Bernoulli
utilisation, 1778
de Rademacher, 1685
intégrable, 1636
suite de variables aléatoire de Bernoulli, 1766
variance, 1637
empirique, 1637, 1711
empirique corrigée, 1711
vecteur gaussien, 1681
variation des constantes, 1449, 1454
vecteur
cyclique, 385
gaussien, 1681
propre, 380
unitaire normal, 1151
unitaire tangent, 1149
Vitali (ensemble), 764
vitesse d'un chemin, 1127
voisinage, 272, 375
volume
d'une région solide, 891
région bornée dans $\mathbb{R}^3$, 888
vraisemblance, 1718

Weiner
constante, 1831
Wronskien, 1482

Liste des notations

$N \lhd G$ Le sous-groupe N est normal dans G, page 107

Algèbre

$\mathcal{C}^1(U, \mathbb{R}^n)$ Les applications une fois continument dérivables, page 625

$\mathcal{L}(E, F)$ Ensemble des applications linéaires de E dans F, page 305

$\mathbb{K}(A)$ corps contenant $\mathbb{K}$ et A, page 186

$\mathbb{K}[A]$ anneau contenant $\mathbb{K}$ et A, page 186

$\mathbb{N}_0$ les naturels non nuls : $\mathbb{N}_0 = \mathbb{N}\backslash\{0\}$, page 92

$\mathcal{M}_{n\times m}$ l'ensemble des matrices $n \times m$, page 309

∇f gradient de la fonction f, page 603

$\mathtt{proj}_V$ projection de $V \times W$ sur V, page 482

$\mathrm{Span}(A)$ l'ensemble des combinaisons linéaires finies d'éléments de A, page 301

$df_a(u)$ Application de la différentielle de f sur le vecteur u, page 591

$f^{(n)}$ La n-ième dérivée de la fonction f, page 634

$o(x)$ fonction tendant rapidement vers zéro, page 635

(p) idéal engendré par p, page 144

$[\mathbb{L} : \mathbb{K}]$ degré d'une extension de corps, page 185

$\mathbb{F}_p$ lorsque p est premier, page 173

$\mathbb{F}_{p^n}$ corps fini à p^n éléments, page 211

$\mathbb{Q}$ corps des fractions sur $\mathbb{Z}$, page 152

$\mathrm{Frac}(\mathbb{A})$ Le corps des fractions de l'anneau $\mathbb{A}$, page 152

$\mathrm{Fun}(X, Y)$ les applications de X vers Y, page 141

$\mathrm{S}(E)$ Les opérateurs auto-adjoints de E, page 446

U_n Le groupe des racines n^{e}de l'unité., page 136

$\mathrm{res}(P, Q)$ résultat des polynômes P et Q, page 353

$\sqrt{A}$ racine d'une matrice hermitienne positive, page 671

$\theta_\alpha(P)$ la multiplicité de α par rapport à P, page 168

$A[X]$ tous les polynômes de degré fini à coefficients dans A, page 165

$A_n[X]$ les polynômes à coefficients dans A et de degré inférieur à n, page 165

$C(P)$ matrice compagnon, page 427

$D \mid P$ D divise P, page 166

$E_\lambda(u)$ Espace propre de u, page 380

$mat_{\mathcal{B}}(q)$ matrice de q dans la base $\mathcal{B}$, page 1108

$U(A)$ ensemble des inversibles, page 142

Ensembles de matrices

$\mathrm{Aut}(E)$ automorphisme de l'espace vectoriel E, page 305

$\mathrm{End}(E)$ les endomorphismes de E, page 305

$L(E,F)$ applications linéaires bornées (continues), page 459

$\mathrm{O}(n,\mathbb{R})$ le groupe des matrices orthogonales, page 364

$\Omega(E)$ formes quadratiques non dégénérées, page 1111

$Q(E)$ formes quadratiques réelles sur E, page 441

$Q^+(E)$ formes quadratiques positives, page 1111

$Q^{++}(E)$ formes quadratiques strictement définies positives, page 1111

$S^+(n,\mathbb{R})$ matrices symétriques définies positives, page 406

$S^{++}(n,\mathbb{R})$ matrices symétriques strictement définies positives, page 406

$S_n^{p,q}(\mathbb{R})$ matrices symétriques réelles de signature (p,q), page 1111

$\beta(s)$ Vecteur unitaire de la binormale, page 1151

$\gamma \sim g$ Équivalence d'arcs paramétrés, page 1142

$\nu(s)$ Vecteur unitaire de la normale principale, page 1151

$c(s)$ rayon de courbure, page 1151

$t(s)$ Torsion, page 1152

Géométrie

$(x_0 : \ldots : x_n)$ coordonnées homogènes dans un espace projectif, page 1202

$\mathcal{C}\mathrm{onv}(A)$ enveloppe convexe, page 500

$\mathbb{C}[G]$ combinaisons d'éléments de G à coefficients dans $\mathbb{C}$, page 1027

$\mathrm{PGL}(E)$ groupe projectif, page 1190

B^o orthogonal dans le dual, page 312

$P(E)$ l'espace projectif de E, page 1183

$P_1(\mathbb{C})$ sphère de Riemann, page 1205

Chaînes de Markov

$\pi(x)$ lié au temps de retour, page 1757

Probabilités et statistique

$\sigma(X)$ La tribu engendrée par la variable aléatoire X, page 1631

$a \wedge b$ $\min(a,b)$, page 1775

K_X matrice de covariance d'un vecteur gaussien, page 1681

$m(\mathcal{A})$ Ensemble des fonctions $\mathcal{A}$-mesurables, page 710

Théorie des groupes

$(G/H)_g$ classes à gauche, page 130

$[a]_p$ ensemble des $a + kp$, page 121

$[G, G]$ groupe dérivé, page 109

$[g, h]$ commutateur dans un groupe, page 109

$\mathrm{Aff}(\mathbb{R}^n)$ Le groupe des applications affines bijectives de $\mathbb{R}^n$., page 508

gr groupe engendré, page 107

$\hat{G}$ groupe des caractères de G, page 1023

σ_x réflexion par rapport à x, page 949

A_n groupe alterné, page 334

$D(G)$ groupe dérivé, page 109

D_n groupe diédral, page 988

G^{ab} groupe abélianisé de G, page 110

$N \times_\phi H$ produit semi-direct, page 134

S_n le groupe symétrique, page 134

Topologie et théorie des ensembles

$\complement A$ Le complémentaire de l'ensemble A, page 90

$\mathrm{Diam}(A)$ Diamètre de la partie A, page 535

∂A La frontière de l'ensemble A, page 378

$A \Delta B$ différence symétrique, page 91

A^c complémentaire de A, page 90

Analyse

$\mathrm{Isom}(X)$ Le groupe des isométries de X, page 248

$\mu \ll \nu$ La mesure μ est absolument continue par rapport à la mesure ν., page 743

$C^\infty\big(\mathbb{R}, \mathscr{S}'(\mathbb{R}^d)\big)$ Fonctions à valeurs dans les distributions., page 1489

$C^\infty(I, \mathscr{D}'(\mathbb{R}^d))$ fonctions à valeurs dans les distributions, page 1417

$(S, \hat{\mathcal{F}}, \hat{\mu})$ complété de l'espace mesuré $(S, \hat{\mathcal{F}}, \hat{\mu})$, page 703

$\mathcal{L}(E, F)$ Les applications linéaires de E vers F, page 613

$\mathcal{L}^{(n)}(V, W)$ L'espace des applications n-linéaires $V^n \to W$, page 629

$\arg(z)$ La valeur principale de l'argument de $z \in \mathbb{C}$, page 1351

L Les applications linéaires continues de E vers F, page 613

$\mathbb{D}_b$ l'ensemble de écritures décimales en base b, page 476

$\mathbb{R}$ l'ensemble des réels, page 101

$\mathbb{R}^+$ les réels positifs ou nuls, page 103

exp exponentielle, page 916

$\mathscr{H}'$ dual, page 1255

$\liminf a_n$ limite inférieure, page 720

$\limsup a_n$ limite supérieure, page 720

$\mathcal{L}^p$ espace de Lebesgue, sans les classes, page 1278

$\mu \perp \nu$ mesures perpendiculaires, page 743

μ^* La mesure extérieure associée à la mesure μ, page 698

∂_z,$\partial_{\bar{z}}$ dérivées partielles d'une fonction complexe, page 1321

$\texttt{proj}_K(x)$ projection orthogonale de x sur y, page 1252

$\sigma(\mathcal{A})$ tribu engendrée par $\mathcal{D}$, page 686

$\mathscr{D}(\Omega)$ Les fonctions C^∞ à support compact sur Ω, page 1404

$\mathscr{S}'(\mathbb{R}^d)$ espace des distributions tempérées, page 1410

$A^2(\Omega)$ espace de Bergman, page 1363

$A^\perp$ orthogonal d'une partie., page 1254

$C_c(I)$ fonctions continues à support compact dans I, page 1293

$f \sim g$ fonctions ayant des limites équivalentes, page 548

$H^1(\Omega)$ espace de Sobolev sur Ω, page 1433

$H^1(I)$ espace de Sobolev, page 1427

$H^m(M)$ espace de Sobolev, page 1434

$L^1_{loc}(I)$ fonctions intégrables sur les compacts de I, page 1428

L^p espace de Lebesgue avec les classes, page 1279

$M_i\varphi$ La fonction $x \mapsto x_i\varphi(x)$, page 1315

$S_n f$ somme partielle de série de Fourier, page 1301

Chapitre 1

Introduction

1.1 Préface

Ce livre s'adresse principalement au candidat à l'agrégation qui y trouvera de quoi dire pour presque toutes les leçons. La matière de ce livre (sous-entendu : tous volumes ensemble) recouvre l'épreuve d'analyse, d'algèbre et de l'option de modélisation.

Petit avertissement : ce livre ne contient pas beaucoup d'exemples, alors que le titre de presque toutes les leçons contient le mot «exemple». Ce livre peut vous aider à

— apprendre la mathématique,
— trouver des développements
— remplir votre plan de théorèmes.

Pour les exemples, il faudra souvent aller chercher ailleurs.

Mon modèle économique

Mon modèle économique est le don. Chaque achat sur thebookedition.com me génère exactement zéro euros. Si vous aimez ce livre, vous pouvez faire un don, surtout si vous avez réussi l'agrégation et que vous pensez que ce livre en est pour quelque chose.

Dons en espèces :

IBAN FR76 3000 4004 0600 0035 2497 784

BIC BNPAFRPPBSC

Dons en nature acceptés : rédaction d'un chapitre, un développement, relecture d'une partie, réponse à l'une ou l'autre des questions qui restent en suspend ...

1.2 Auteurs, contributeurs, sources et remerciements

Les remerciements, dans chaque catégorie, sont mis dans l'ordre chronologique approximatif. Les noms en couvertures sont ceux qui ont fourni du code LaTeX (typiquement : un patch via github), par ordre chronologique approximatif d'entrée dans le projet.

1.2.1 Ceux qui ont travaillé sur le Frido

(1) Carlotta Donadello pour l'ensemble du cours de CTU de géométrie analytique 2010-2011. Une grosse partie de «analyse réelle» vient de là.

(2) Les étudiants de géométrie analytique en CTU 2010-2011 ont détecté d'innombrables coquilles. Les étudiants du cours présentiel de géométrie analytique 2011-2012 ont signalé un certain nombre d'incorrections dans les exercices et les corrigés. Les agrégatifs de Besançon 2011-2012 pour leurs plans et leurs développements.

(3) Lilian Besson pour m'avoir signalé un paquet de fautes, et quelques points pas clairs en statistiques.

(4) Plouf qui m'a signalé une coquille dans le fil la-selection-scientifique-de-la-semaine-numero-106.

(5) Benjamin de block pour des coquilles et une mise au point sur les conventions à propos de $\mathbb{R}^+$ et $(\mathbb{R}^+)^*$.

(6) Olivier Garet pour avoir répondu à plein de questions de probabilités.

(7) François Gannaz pour de la relecture et une version plus claire de la preuve (et de l'énoncé) de la proposition 3.101.

(8) Danarmk pour des réponses à des questions `http://linuxfr.org/nodes/110155/comments/1675589`.

(9) Cédric Boutilier pour des réponses à des questions de probabilité statistique. `https://github.com/LaurentClaessens/mazhe/issues/16`

(10) Remsirems pour des réponses à des questions d'analyse. `http://linuxfr.org/nodes/110155/comments/1675813`

(11) Bertrand Desmons pour plusieurs patch rendant plus clairs de nombreux passages sur les suites de Cauchy dans $\mathbb{Q}$.

(12) Anthony Ollivier pour m'avoir fait remarquer qu'il n'est pas vrai que $A[X]$ est euclidien lorsque A est intègre (contre-exemple : $A = \mathbb{Z}$). Ça fait une faute de moins dans le Frido.

1.2.2 Aide directe, mais pas volontairement sur le Frido

(1) Plein de monde pour diverses contributions à des énoncés d'exercices. Pierre Bieliavsky pour les énoncés d'analyse numérique (MAT1151 à Louvain la Neuve 2009-2010). Jonathan Di Cosmo pour certaines corrections de MAT1151. François Lemeux, exercices sur les normes de matrices et correction de coquilles. Martin Meyer et Mustapha Mokhtar-Kharroubi pour certains exercices du cours *Outils mathématiques* (surtout ceux des DS et examens).

(2) Nicolas Richard et Ivik Swan pour les parties des exercices et rappels de calcul différentiel et intégral (Université libre de Bruxelles, 2003-2004) qui leurs reviennent.

(3) Carlotta Donadello pour la partie géométrie analytique : topologie dans $\mathbb{R}^n$, courbes, intégrales, limites. (Université de Franche-Comté 2010-2012)

(4) Le forum usenet de math, en particulier pour la construction des corps fini dans la fil « Vérifier qu'un polynôme est primitif » initié le 20 décembre 2011.

(5) Mihai Bostan nous a donné ses notes manuscrites de son cours présentiel de géométrie analytique 2009-2010. (Presque) Toute la structure du cours de géométrie analytique lui est due (qui est maintenant fondue un peu partout dans les chapitres d'analyse).

1.2.3 Des gens qui ont fait un travail qui m'a bien servi

(1) Arnaud Girand pour avoir mis ses développements bien faits en ligne. Une bonne vingtaine de résultats un peu partout dans ces notes viennent de lui.

(2) Le site `http://www.les-mathematiques.net` m'a donné les preuves de nombreux résultats.

(3) Pierre Monmarché pour son document en ligne tout plein de développements, et des réponses à des questions.

(4) Tous les contributeurs du Wikipédia francophone (et aussi un peu l'anglophone) doivent être remerciés. J'en ai pompé des quantités astronomiques ; des articles utilisés sont cités à divers endroits du texte, mais ce n'est absolument pas exhaustif. Il n'y a à peu près pas un résultat important de ces notes dont je n'aie pas lu la page Wikipédia, et souvent plusieurs pages connexes.

(5) Les intervenants du fil «Antisymétrisation, alterné, déterminant et caractéristique» sur `les-mathematiques.net` m'ont bien aidé pour la section sur les déterminants 9.1 (bien qu'ils ne le savent pas).

(6) Xavier Mauquoy pour l'énoncé et la preuve du théorème 12.24.

(7) David Revoy pour les dessins de Pepper&Carrot de la couverture.

J'ai souvent donné entre parenthèse à côté des mots « théorème », « lemme » ou « proposition » une ou plusieurs références vers les sources de la preuve que je donne. Ce sont parfois des liens vers la bibliographie ; parfois aussi des liens hypertexte vers des sites, des blogs, etc. Tous ces gens ont fait du bon boulot. Sans toute cette « communauté », l'internet serait mort [1].

1.3 Les choses qui doivent vous faire tiquer

Un cours de math doit toujours être lu attentivement, surtout si vous avez l'intention de resservir à un jury le fruit de vos lectures. Dans ce livre, trois éléments doivent vous faire redoubler de prudent.

D'abord les références à [1] indiquent qu'une bonne partie de ce qui suit est de l'invention personnelle de l'auteur. Ensuite les notes en bas de pages écrites en `fonte tt` comme celle-ci [2] Et enfin certains problèmes sont indiqués plus longuement dans un environnement dédié en petit caractères comme ceci : **Problèmes et choses à faire**

Les choses écrites comme ceci sont des questions ou des éléments sur lesquels j'ai un doute. Lisez-les attentivement. Ces notes mentionnent des points que personnellement je n'oserais pas affirmer plein d'aplomb à un jury d'agrégation.

1.4 Les questions pour lesquelles je n'ai pas (encore) de réponse

Ces notes ne sont pas relues de façon systématique. Aucune garantie. Merci de me signaler toute faute ou remarque : le relecteur c'est toi. Voici une petite liste de questions que je me pose ou de choses écrites dont je ne suis pas certain. Si vous avez un avis ou une réponse à un des points, merci de vous faire connaître.

Vous pouvez soit m'écrire directement, soit créer une *issue* sur github pour répondre aux questions.

1.4.1 Mes questions d'algèbre, géométrie.

(1) Est-ce qu'il existe une structure raisonnable d'espace vectoriel sur $\mathbb{Z}$? Est-ce qu'il existe des corps discrets infinis ?

Dans cette question, j'ai derrière la tête que dans un espace vectoriel topologique nous avons une notion de suite de Cauchy, définition 6.37. Donc dans ce cas la notion d'espace complet et une notion topologique. Or il y a l'exemple 6.42 qui donne deux distances sur $\mathbb{N}$, qui donnent la même topologique, mais l'un étant complet, l'autre non.

Si il y avait une structure vectorielle sur $\mathbb{N}$, cela créerait une contradiction. Au moins au sens où la définition 6.37 de suite de Cauchy «topologique» ne redonne pas la même que la notion «métrique» de la définition usuelle 6.39.

(2) La «décomposition en facteurs premiers» dans $\mathbb{Z}[i\sqrt{2}]$ que je donne dans l'exemple 4.87 est-elle correcte ? En particulier le lemme 4.88 ?

(3) Est-ce que la fin de la démonstration 5.124 avec cette histoire d'ensemble $\{\xi_k^q$ tel que $q \in \mathbb{N}\}$ fini est compréhensible ?

(4) Les représentations *irréductibles* sont les modules *indécomposables*. Quid des modules irréductibles ? C'est pas un peu dingue de ne pas utiliser le mot «irréductible» pour désigner les mêmes choses dans le cas des modules et celui des représentations ?

(5) Rendre rigoureuse la remarque (9.219) qui dit que les matrices ont le polynôme minimal est égal au polynôme caractéristique sont denses dans les matrices.

1. Cette dernière phrase doit être comprise comme un appel à ne pas utiliser Moodle et autres iCampus pour diffuser vos cours de math, ou en tout cas pas comme moyen exclusif.

2. Les notes comme celle-ci signifient qu'il y a quelque chose dont je ne suis pas sûr.

(6) Soit $\mathbb{L}$ une extension algébrique de corps de $\mathbb{K}$ et $a \in \mathbb{L}$. Est-ce que le polynôme minimal de a dans $\mathbb{K}[X]$ est l'unique polynôme irréductible unitaire $P \in \mathbb{K}[X]$ tel que $P(a) = 0$? D'ailleurs, est-ce qu'un tel polynôme existe ? est unique ? J'utilise cela dans la proposition 5.92.

(7) La partie initiation de récurrence ($r = 2$) de la preuve de la proposition 11.33 à propos de convexe et de barycentre est-elle correcte ? Ce passage de l'espace affine à l'espace vectoriel sous-jacent me paraît un peu facile.

(8) Le lemme 5.43 à propos de PGCD de polynômes est-il correct ? J'imagine que le polynôme $\operatorname{pgcd}(P, PU + R)$ ne dépend pas de U. Est-ce exact ?

(9) Est-ce que parler de sous-groupe «normal» au lieu de «distingué» est un anglicisme ?

(10) Est-ce que l'énoncé et la démonstration de la proposition 5.63 sont corrects ? Si a et b sont des racines de P, alors $\mu_a\mu_b$ divise P (si $\mu_a \neq \mu_b$). Cette proposition est utilisée dans la démonstration de l'irréductibilité des polynômes cyclotomiques (proposition 5.122).

(11) À quoi sert l'hypothèse «autre que $\mathbb{F}_2$» dans le lemme 9.4 ? Peut-être dans la notion de déterminant parce qu'en caractéristique 2, l'antisymétrie d'une forme linéaire n'implique le fait qu'elle soit alternée.

(12) L'inversibilité de la somme de Gauss (proposition 5.138) est-elle bien démontrée ?

(13) Des commentaires sur l'exemple 5.80 qui montre que $X^p - X + 1$ est irréductible sur $\mathbb{F}_p$.

(14) Les idéaux de A/I sont en bijection avec les idéaux de A contenant I. Justification de l'équation (4.17).

(15) L'énoncé et la démonstration d'une des multiples version du théorème de l'élément primitif, proposition 5.163.

(16) Pourquoi la pseudo-réduction simultanée (corollaire 9.173) est-elle *pseudo* ? Pourtant les matrices sont bel et bien simultanément diagonalisées.

(17) À propos d'extensions algébriques, est-ce que la proposition 5.75 est correcte ? Est-ce qu'implicitement, il n'y a pas un sur-corps de $\mathbb{K}$ dans lequel il faut travailler ?

(18) Est-ce que la proposition 5.50 disant qu'un polynôme minimal est irréductible et premier avec tout polynôme non annulateur est correcte ?

(19) À propos de construction à la règle et au compas. Pour l'addition d'angles, l'exemple 5.176 explique comment on construit la somme de deux angles. Le problème est que cette construction se fait par intersection de deux cercles. Une des deux intersections donne $\alpha+\beta$ et l'autre donne $\alpha - \beta$. Comment par construction peut-on choisir le bon point ?

(20) À propos de chiffrement RSA, quelle est la probabilité que le message M ne soit pas premier avec p ? Est-ce que Alice (qui est celle qui chiffre avec la clef de Bob) peut le vérifier ? Que penser des points que j'énumère à la page 465 au dessus du problème 13 ?

(21) Extension des scalaires. Si $\{e_i\}_i$ est une base du $\mathbb{K}$-espace vectoriel E, alors $\{1 \otimes e_i\}$ est une base de $E_{\mathbb{L}} = \mathbb{L} \otimes_{\mathbb{L}} E$ où $\mathbb{L}$ est une extension de KK. C'est vrai ? Que penser de la preuve donné au lemme 9.204 ?

(22) Si $f\colon V \to V$ est une application linéaire, on ne peut pas définir la transposée $f^t\colon V \to V$ comme étant l'application linéaire dont la matrice est la transposée de la matrice de f parce que cette définition n'est pas invariante par changement de base. C'est vrai ? Voir le problème 9.

(23) En géométrie projective, dans la sphère de Riemann $\hat{\mathbb{C}} = P_1(\mathbb{C}) = \mathbb{C} \cup \{\infty\}$ est-ce qu'il existe une notion de cercle dont le centre est ∞ ? Voir le point 21.65.

(24) Géométrie projective. Tout le monde semble définir le birapport en identifiant $P(\mathbb{K}^2)$ à $\hat{\mathbb{K}} = \mathbb{K} \cup \{\infty\}$. Bien entendu, personne ne semble s'être attribué la mission d'expliciter la dépendance du birapport en le choix de l'identification. Je le fais à la définition 21.40.

Mais cette définition dépend du choix d'identification $\varphi\colon P(\mathbb{K}^2) \to \hat{\mathbb{K}}$, comme le montre l'exemple 21.42. J'ai donc défini des classes d'identifications possibles $A(\varphi)$ en 21.38. Et je

démontre la proposition 21.43 que si $\varphi_a \in A(\varphi)$ alors les birapports construits à partir de φ et φ_a sont identiques.

Question : pourquoi personne ne semble faire ce travail ? En quoi l'identification φ_0 que tout le monde utilise est plus canonique qu'une autre ? Est-ce que l'on peut décrire simplement les classes $A(\varphi)$? Le groupe qui conserve le birapport associé à φ est-il isomorphe au groupe qui conserve le birapport associé à φ' ? Quels que soient φ et φ' ?

Suis-je la seule personne au monde à m'être demandé si le birapport était un objet canonique ?

(25) En géométrie projective, dans $P_1(\mathbb{C}) = \hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, si ℓ est une droite dans $\mathbb{C}$, est-ce que la droite correspondante dans $\hat{\mathbb{C}}$ contient le point ∞ ? Moi j'ai envie de dire que ∞ est sur toutes les droites. Voir le problème 23 et la remarque 21.63.

(26) Géométrie affine, barycentre. Les mauvaise langues diraient que tout cela est du snobisme autour de la paresse d'écrire $\overrightarrow{xy}$ au lieu de $y - x$. Est-ce qu'il y a des cas où toute l'étude des espaces affines et des barycentres en particulier apportent *réellement* plus qu'une facilité d'écriture par rapport à travailler dans le cadre vectoriel pur ?

(27) Isomorphisme du corps $\mathbb{R}$. Que penser de la remarque 5.5 ?

1.4.2 Mes questions de probabilité et statistiques.

(1) Est-ce que le théorème d'arrêt de Doob 37.14 est correctement énoncé ? En particulier la seconde condition.

(2) À propos du problème de la ruine du joueur. Dans [2], l'équation (37.66) vient avec une lim sup et non une limite normale. Je ne comprends pas pourquoi.

(3) Soit une variable aléatoire X à valeurs réelles. Est-ce que la tribu engendrée par X est d'une façon ou d'une autre engendrée par les «courbes de niveau» de X ? C'est à dire par les $X^{-1}\big(\{t\}\big)$ pour les $t \in \mathbb{R}$.

C'est ce qui semble ressortir de l'exemple de 34.2.10. Et intuitivement, je trouve que ça irait bien ...

1.4.3 Les preuves à relire par des experts

Ici lorsque je parle d'«expert», j'entends quelqu'un qui est sûr de son coup dans un domaine particulier. Soit quelqu'un qui donne un cours dans le domaine, soit un étudiant qui a eu une bonne note dans un cours du domaine. Je m'adresse donc à des personnes qui ont déjà lu et compris une preuve des résultats *avant* de les lire ici.

Les preuves des résultats suivants sont en tout ou en partie des inventions personnelles, et doivent donc être lues avec beaucoup de prudence. Si vous les lisez, que vous ne voyez pas de fautes et que vous êtes sûr de votre coup, écrivez moi. Si au contraire, vous y trouvez des fautes ou des raccourcis, écrivez moi même si vous n'êtes pas sûr de votre coup.

(1) La proposition 14.72.

Proposition 1.1 (Dérivation sous l'intégrale)**.**
Supposons $A \subset \mathbb{R}^m$ ouvert et $B \subset \mathbb{R}^n$ compact. Nous considérons une fonction $f : A \times B \to \mathbb{R}$. Si pour un $i \in \{i, \dots, n\}$, la dérivée partielle $\frac{\partial f}{\partial x_i}$ existe dans $A \times B$ et est continue, alors la fonction

$$F(x) = \int_B f(x,t)dt \tag{1.1}$$

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

$$\frac{\partial F}{\partial x_i}(a) = \int_B \frac{\partial f}{\partial x_i}(a,t)dt, \tag{1.2}$$

pour tout a dans l'ouvert A.

(2) La proposition 13.137

Proposition 1.2.
Soit un espace mesuré $(S, \mathcal{F}, \mu)$ et une fonction mesurable positive $w\colon S \to \bar{\mathbb{R}}^+$. Alors la formule

$$(w \cdot \mu)(A) = \int_A w d\mu \tag{1.3}$$

*pour tout $A \in \mathcal{F}$ définit une mesure positive sur $(S, \mathcal{F})$ appelée **produit** de la mesure μ par la fonction w. La fonction w est la **densité** de la mesure $w \cdot \mu$ par rapport à la mesure μ.*

(3) La proposition 9.207 et la remarque 9.208.
Soit $\{e_i\}$ une base de E et $\{e_a\}$ une de$F$. Nous noterons également e_i et e_a les éléments τe_i et τe_a correspondants. Si la fonction $f\colon E_{\mathbb{L}} \to F_{\mathbb{L}}$ s'écrit dans ce ces bases comme

$$f(e_i) = \sum_a f_{ai} e_a \tag{1.4}$$

alors nous définissons $\texttt{proj}(f)$ par

$$(\texttt{proj}\, f)e_i = \sum_a \texttt{proj}(f_{ai})e_a. \tag{1.5}$$

Proposition 1.3.
L'application `proj` définie en (1.5) *est indépendante du choix des bases.*

(4) La proposition 9.215

Proposition 1.4.
Soit $\mathbb{L}$ une extension du corps $\mathbb{K}$ et une application linéaire $f\colon E \to F$ entre deux $\mathbb{K}$-espaces vectoriels. Alors $\mu_f = \mu_{f_{\mathbb{L}}}$.

(5) Le lemme 13.27

Lemme 1.5.
Résultats sur les unions croissantes d'ensembles mesurables dans $(S, \mathcal{A}, \mu)$.

(a) Si (A_k) est une suite croissante d'ensembles μ-mesurables dont l'union est mesurable, alors

$$\lim_{n\to\infty} \mu(A_k) = \mu(\bigcup_k A_k). \tag{1.6}$$

(b) Soit K_n, une suite emboîtée d'éléments de $\mathcal{A}$ tels que $K_n \to S$. Si $A \in \mathcal{A}$ alors

$$\lim_{n\to\infty} \mu(A \cap K_n) = \mu(A). \tag{1.7}$$

(6) Le lemme 13.104

Lemme 1.6.
Soit $f\colon (S, \mathcal{F}) \to \bar{\mathbb{R}}$ une fonction positive mesurable. Il existe une suite $f_n\colon S \to \mathbb{R}$ de fonctions étagées positives telles que $f_n \to f$ ponctuellement et $f_n \leqslant f$.

(7) La proposition 2.33

Proposition 1.7.
Soit la suite de rationnels (x_n) définie par $x_0 \in \mathbb{Q}^+$ et

$$x_{n+1} = x_n + \frac{x_n^2 - 2}{2x_n}. \tag{1.8}$$

Alors :

(a) *En posant* $y_n = x_n^2$ *nous avons* $y_n \to 2$.

(b) *La suite* (x_n) *est de Cauchy.*

(c) *La suite* (x_n) *ne converge pas dans* $\mathbb{Q}$.

(8) La proposition 19.83

Proposition-définition 1.8 (Cercle circonscrit)**.**
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 :

(a) *Le cercle circonscrit à* γ *coupe* Γ *en au moins deux points distincts.*

(b) *Tout arc du cercle circonscrit plus grand que le demi-cercle intersection* Γ.

(9) La proposition 3.102

Proposition 1.9.
Soient des entiers positifs $\alpha_1, \ldots, \alpha_p$. *Nous avons*

$$\bigcup_{i=1}^{p} U_{\alpha_i} = \{1\} \tag{1.9}$$

si et seulement si $\operatorname{pgcd}(\alpha_1, \ldots, \alpha_p) = 1$ *(c'est à dire que les* α_i *sont premiers dans leur ensemble).*

(10) La proposition 5.50

Proposition 1.10.
Soit $\mathbb{L}$ *un extension de* $\mathbb{K}$ *et* $a \in \mathbb{L}$ *dont le polynôme minimal sur* $\mathbb{K}$ *est* $\mu_a \in \mathbb{K}[X]$. *Alors*

(a) *le polynôme* μ_a *est irréductible sur* $\mathbb{K}$ *;*

(b) *si* $Q \in \mathbb{K}[X]$ *est un polynôme non annulateur de* a *alors* Q *et* μ_a *sont premiers entre eux.*

(11) La proposition 5.75

Proposition 1.11 (Propriétés d'extensions algébriques[1])**.**
Soit $\mathbb{K}$ *un corps commutatif et* a *un élément algébrique sur* $\mathbb{K}$, *de polynôme minimal* μ_a *de degré* n. *Alors*

(a) *En considérant l'application d'évaluation*

$$\begin{aligned} \varphi_a\colon \mathbb{K}[X] &\to \mathbb{L} \\ Q &\mapsto Q(a), \end{aligned} \tag{1.10}$$

nous avons $\mathbb{K}[a] = \operatorname{Image}(\varphi_a)$.

(b) *Une base de* $\mathbb{K}[a]$ *comme espace vectoriel sur* $\mathbb{K}$ *est donnée par* $\{1, a, a^2, \ldots, a^{n-1}\}$.

(c) *Le degré de l'extension* $\mathbb{K}[a]$ *est égal au degré du polynôme minimal :*

$$\big[\mathbb{K}[a] : \mathbb{K}\big] = n. \tag{1.11}$$

(d) *L'anneau* $\mathbb{K}[a]$ *est l'ensemble des polynômes en* a *de degré* $n-1$ *à coefficient dans* $\mathbb{K}$.

(e) $\mathbb{K}(a) = \mathbb{K}[a]$.

(f) $\mathbb{K}[a] \simeq \mathbb{K}[X]/(\mu_a)$ *(isomorphisme d'anneau).*

(12) La proposition 6.59

Proposition 1.12 ([1])**.**
Quel que soit le réel r, il existe une suite croissante de rationnels convergente vers r.

(13) Le lemme 6.83

Lemme 1.13.
Soit (X, τ_X) un espace topologique et $S \subset X$, un fermé de X sur lequel nous considérons la topologie induite τ_S. Si F est un fermé de (S, τ_S) alors F est fermé de (X, τ_X).

(14) Le lemme 6.122

Lemme 1.14.
Si K_1 et K_2 sont des compacts dans $\mathbb{R}$ alors $K_1 \times K_2$ est compact dans $\mathbb{R}^2$.

(15) La proposition 6.138

Proposition 1.15 (Caractérisation séquentielle de la continuité)**.**
Le lien entre continuité et continuité séquentielle.

(a) Une application continue est séquentiellement continue (quels que soient les espaces de départ et d'arrivée).

(b) Si X et Y sont des espaces métriques, alors une fonction $f\colon X \to Y$ est continue en un point si et seulement si elle est y séquentiellement continue.

(16) Le théorème 6.162

Théorème 1.16.
Un produit fini d'espaces métriques non vides est compact si et seulement si chacun de ses facteurs est compact.

(17) Le lemme 16.108

Lemme 1.17.
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 $k = 0, \ldots, n-1$.

(18) La proposition 12.88

Proposition 1.18.
Si la fonction réelle $f\colon I = [a, b[\to \mathbb{R}$ est croissante et bornée, alors la limite

$$\lim_{x \to b} f(x) \tag{1.12}$$

existe et est finie.

(19) La proposition 18.79

Proposition 1.19.
Une fonction dérivable sur un intervalle I de $\mathbb{R}$

(a) est convexe si et seulement si sa dérivée est croissante sur I.

(b) est strictement convexe si et seulement si sa dérivée est strictement croissante sur I

(20) La proposition 18.85

Proposition 1.20.
Soit $f\colon \mathbb{R} \to \mathbb{R}$ une fonction convexe et $a \in \mathbb{R}$. Il existe une constante $c_a \in \mathbb{R}$ telle que pour tout x nous ayons

$$f(x) - f(a) \geqslant c_a(x - a). \tag{1.13}$$

Autrement dit, le graphe de la fonction f est toujours au dessus de la droite d'équation

$$y = f(a) + c_a(x - a). \tag{1.14}$$

Je me demande si ce ne serait pas un «si et seulement si».

(21) La proposition 18.86

Proposition 1.21.
Si g est une fonction convexe, il existe deux suites réelles (a_n) et (b_n) telles que

$$g(x) = \sup_{n\in\mathbb{N}}(a_n x + b_n). \tag{1.15}$$

(22) Le théorème 13.203

Théorème 1.22.
Soit Q un compact connexe par arcs et une fonction continue $f\colon Q \to \mathbb{R}$. Si λ est la mesure de Lebesgue, alors il existe $a \in Q$ tel que

$$f(a) = \frac{1}{\lambda(Q)} \int_Q f d\lambda \tag{1.16}$$

(23) Le lemme 19.85

Lemme 1.23.
Soit une courbe simple, 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 Γ est localement le graphe d'une fonction convexe.

(24) La proposition 19.86.

Proposition 1.24.
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 $\partial D = \Gamma$.

(25) La proposition 15.39

Proposition 1.25.
Si la série entière $\sum_{n\geqslant 0} a_n z^n$ a un rayon de convergence R alors

(a) La somme est une fonction holomorphe dans le disque de convergence.

(b) La somme est différentiable et

$$du_{z_0}(z) = \sum_{n=1}^{\infty} n a_n z_0^{n-1} z. \tag{1.17}$$

(c) De plus pour tout $z_0 \in B(0,R)$, on pose[3]

$$S(z) = \sum_{n\geqslant 0} a_n z^n \tag{1.18a}$$

$$T(z) = \sum_{n\geqslant 1} n a_n z^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} z^n. \tag{1.18b}$$

Alors nous avons

$$\lim_{z\to z_0} \frac{S(z) - S(z_0)}{z - z_0} = T(z_0). \tag{1.19}$$

(26) Le corollaire 15.43

3. Pour rappel, dans tout ce texte, $B(a,r)$ est une boule *ouverte*.

Corollaire 1.26.
La somme d'une série entière est de classe C^∞ sur le disque ouvert de convergence.

(27) La proposition 12.37

Proposition 1.27.
Soit $f\colon \mathbb{R}^2 \to \mathbb{R}$ une application continue dont la variable y varie dans un compact I de $\mathbb{R}$. Alors la fonction

$$\begin{aligned} d\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \sup_{y\in I} f(x,y) \end{aligned} \tag{1.20}$$

est continue.

À ce propos, est-il vrai que l'opération sup vue comme application $C(\mathbb{R}^n) \to C(\mathbb{R}^{n-1})$ est continue ? Pour quelle topologie sur les espaces de fonctions continues ?

(28) La proposition 19.93

Proposition 1.28.
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

$$\gamma\big([s_p, s_q]\big) = [p,q] \tag{1.21}$$

soit

$$\gamma\big([s_q, L]\big) = [p,q] \tag{1.22}$$

(29) Le corollaire 24.79

Corollaire 1.29.
Soit φ une fonction Schwartz sur $\mathbb{R}^m \times \mathbb{R}^n$. Alors la fonction

$$y \mapsto \sup_{x\in\mathbb{R}^n} |\varphi(x,y)| \tag{1.23}$$

est intégrable.

(30) La proposition 27.13

Proposition 1.30.
Soit $\varphi \in \mathscr{S}(\mathbb{R}^n \times \mathbb{R}^m)$ et la transformée de Fourier partielle

$$\tilde{\varphi}(x,k) = \int_{\mathbb{R}^m} e^{-iky}\varphi(x,y)dy. \tag{1.24}$$

Alors $\tilde{\varphi} \in \mathscr{S}(\mathbb{R}^n \times \mathbb{R}^m)$.

Cette proposition utilise le corollaire 1.29.

(31) Le théorème du flot d'une équation différentielle
Le théorème 30.28 qui donne la régularité C^1 du flot d'une équation de transport est à relire attentivement. Il y a d'inextricables problèmes de notations dont j'espère m'être extirpé, mais sans garanties.

1.5 Comment contribuer et aider (math) ?

1.5.1 Niveau facile

Niveau facile : un étudiant de licence devrait pouvoir le faire.

(1) M'écrire pour me signaler toutes les fautes que vous voyez, même si vous n'êtes pas sûr.

(2) Si vous n'êtes pas expert, me signaler tous le endroits qui vous semblent obscurs. Vu que ces notes sont destinées à *appendre*, les avis des non experts sont très importants.

(3) Mettre une copie de (ou un lien vers) ces notes sur votre site.

1.5.1.1 Géométrie

(1) À propos de représentation de groupe et de triangle équilatéral, l'exemple 17.7 n'est pas terminé. Il y a une phrase inachevée et je ne sais plus ce que je voulais dire. Deviner la suite.

(2) Donner une interprétation en termes de plans, droites ou je ne sais quoi de géométrique au 4-cycle parmi les permutations des sommets du tétraèdre. D'où sort géométriquement la matrice (17.117) ?

(3) La définition 9.121 du polynôme caractéristique est $\chi_u(\lambda) = \det(u - \lambda\mathbb{1})$. Je crains qu'il y ait des incohérences et que la définition prises à certains endroits est $\det(\lambda\mathbb{1} - u)$ (comme sur Wikipédia). Traquer les utilisations de polynôme caractéristique et me signaler les problèmes éventuels (ou les corriger directement).

1.5.2 Niveau moyen

Niveau moyen : un candidat à l'agrégation de mathématique devrait pouvoir le faire.

(1) Donner une liste de résultats qui devraient être connue pour l'agrégation. En particulier, remplir le fichier `98_liste_developpements`

(2) Mettre de l'ordre dans les énoncés des développements de Taylor. En particulier, donner des énoncés corrects pour des fonctions $\mathbb{R}^n \to \mathbb{R}^m$. Faire une référence vers le résultat qu'il faut là où se trouve la référence `382218354`, proposition 18.68(1).

(3) Réfléchir au problème suivant : comment trouver plus facilement les résultats dans le Frido ? Pour l'instant il y a l'index (qui à mon avis ne sert pas à grand chose), l'index thématique et la table des matières. Comment réorganiser ?

Exemple. Je sais qu'il y a un résultat quelque part qui dit que le problème $Ax = b$ est bien posé si le conditionnement de la matrice A est petit. Comment trouver l'énoncé exact ?

1.5.2.1 Questions d'analyse

(1) Est-ce que l'énoncé du théorème de Müntz 18.16 est correct ? Voir en particulier la remarque 18.17 à propos de la présence du monôme 1 dans la liste.

(2) À propos de formule sommatoire de Poisson, est-ce que l'exemple 27.10 est bien fait ? En particulier la formule (27.44) est-elle correcte et bien justifiée ?

(3) Que penser de la remarque 23.39 qui dit qu'on doit avoir un théorème de complétion de partie orthonormale en une base orthonormale pour un espace de Hilbert ? C'est vrai ?

(4) Prouver la proposition 23.12 sur le supplémentaire topologique.

(5) À propos de différentiabilité pour une application entre deux espaces de Banach. Le théorème 12.231 dit que si

$$\frac{d}{dt}\Big[f(x+tu)\Big]_{t=0} \tag{1.25}$$

existe pour tout $x \in B(a,r)$ et est continue (par rapport à x) en $x = a$, et que de plus $\frac{\partial f}{\partial u}(a) = 0$ pour tout u, alors f est différentiable en a.

Est-ce que cet énoncé est encore vrai lorsque $\frac{\partial f}{\partial u}(a) \neq 0$? Est-ce qu'il existe un théorème comparable à la proposition 12.204 pour la dimension infinie ? Il me semble que non, mais je n'ai pas de contre-exemples en tête. Par contre, je vois bien où la preuve bloque : c'est le lien entre la linéarité et la différentiabilité.

(6) Pour démontrer qu'une série entière est de classe C^∞ dans son disque de convergence, on s'appuie sur la proposition 15.39 qui donne la différentiabilité et ensuite sur le corollaire 15.43 qui effectue une parieuse récurrence. Toutes ces preuves sont de moi. À relire.

(7) Faire le lien entre les fonctions trigonométriques (définies par leur série) et les relations dans un triangle serait une bonne affaire. Indice : la relation entre les fonctions trigonométriques et les rotations de $\mathbb{R}^2$ est dans le lemme 16.113 qui a besoin de quelque identités trigonométriques, données dans le lemme 16.23.

1.5.2.2 Questions de géométrie

(1) Démontrer la proposition 11.3 qui donne les relations de Chasles pour un espace affine.

1.5.2.3 Questions de numérique

(1) L'erreur de cancellation provoquée par la différence $a - \tilde{a}$ lorsque a et $\tilde{a}$ n'a pas de conséquences sur l'ordre de grandeur de la réponse. Seulement des conséquences sur la valeur des chiffres significatifs. Vrai ou faux ?

Voir la remarque 32.20.

1.5.3 Niveau difficile

Demande probablement de connaissances avancées en mathématique ; au moins être tout à fait à l'aise avec le niveau d'agrégation.

(1) Répondre aux choses dites «problème et choses à faire» un peu partout.

(2) Lire les rapports du jury et tenir compte des remarques. Il est souvent écrit «on peut faire le lien avec ...» ou «ceci fournit de bons exemples de ...». Donner ces liens (au moins énoncer des théorèmes), écrire du texte expliquant les liens suggérés par le jury.

(3) Un peu partout des `%TODO` sont placés dans les sources LaTeX. Ils décrivent des choses qu'il serait bon de faire. Si vous avez un avis sur l'un d'eux, n'hésitez pas à me le faire savoir.

(4) Préciser l'énoncé et donner une démonstration de la proposition 9.186 qui traite de sommes dénombrables.

(5) Il me faudrait un exemple de partie de $\mathbb{R}$ qui soit non borélienne mais mesurable au sens de Lebesgue. Comme je suis exigeant, je le veux le plus constructif possible : pas d'axiome du choix. Pour l'instant le mieux que j'aie trouvé est une feuille de TD[3] qui utilise des ordinaux.

(6) Est-ce que la proposition 6.52 qui donne le critère $d(x_p, x_q) \leqslant \epsilon$ pour être une suite de Cauchy est valide dans un espace topologique métrique au lieu de normé ? Dans quels cas a-t-on

$$d(a, b) = d(a + u) + d(b + u) \tag{1.26}$$

lorsque d est une distance qui n'est pas spécialement induite d'une norme ?

(7) Pour le théorème de Lax-Milgram 23.56, il y a la majoration $\|u\| \leqslant MC/\alpha$, mais je ne parviens pas à la prouver : j'obtiens seulement $\|u\| \leqslant C/\alpha$. Ici, α est la constante de coercivité, M est tel que $|a(u, v)| \leqslant M\|u\|\|v\|$ et C est la norme de L.

(8) Explosion en temps fini. C'est le corollaire 30.18. Si nous sommes dans le cas où $\lim_{t\to t_{max}} \|y(t)\| = \infty$ alors la dérivée de y n'est pas non plus bornée. Correct ?

1.5.3.1 Questions d'analyse

(1) À propos de suites de Cauchy dans un espace vectoriel topologique et dans un espace métrique, est-ce que le théorème 6.49 est correct ?

Soit V un espace vectoriel topologique métrisable [4], alors il admet une métrique d compatible avec la topologie telle que une suite dans V est d-Cauchy si et seulement si elle est τ-Cauchy.

Dans cet ordre d'idée, il faut des exemples de :

— un espace vectoriel topologique métrisable et une métrique d compatible avec la topologie, mais dont les suites d-Cauchy ne sont pas celles τ-Cauchy. Et en particulier dont la complétude est différente que celle de la «bonne» métrique donnée par la théorème 6.47.

— Et aussi un exemple pour la remarque 6.50.

(2) L'exemple 14.70 parle d'inverser une intégrale et une dérivée au sens des distributions pour prouver que la dérivée de $\int_0^x g(t)dt$ par rapport à x est g. Rendre cela rigoureux.

(3) À propos du théorème de récurrence de Poincaré 13.81, l'application ϕ doit être mesurable ? Répondre à la question posée sur la page de discussion de l'article sur wikipédia.

Toujours à propos du théorème de récurrence de Poincaré, il me semble qu'il y a un énoncé qui insiste sur la compacité de l'espace des phase et une démonstration utilisant la propriété de sous-recouvrement fini. Je serais content de retrouver cela. (ce serait sans doute mettable dans la leçon sur l'utilisation de la compacité)

(4) Dans [4], on parle de la proposition 28.34 à sa page 10. Comment est-ce qu'on justifie le passage

$$\int_{\mathbb{R}^d} T\big(y \mapsto \varphi(x)\psi(x-y)\big)dx = T\Big(y \mapsto \int_{\mathbb{R}^d} \varphi(x)\psi(x-y)dx\Big). \tag{1.27}$$

Sylvie Benzoni précise que «ceci demanderai quelque justification». Où trouver lesdites justifications ? Il s'agit de permuter une distribution et une intégrale.

(5) En ce qui concerne les questions «fines» de topologie concernant l'espace $\mathscr{D}(K)$ des fonctions C^∞ à support dans le compact K, il faut voir les énoncés, les preuves et l'enchaînement des théorèmes et propositions

(a) Proposition 28.7 pour dire que $\mathscr{D}(K)$ est métrique et complet (et donc de Baire).

(b) La proposition 12.299 qui donne la complétude de $\big(C(X,Y), \|.\|_\infty\big)$ lorsque X est compact et Y métrique complet.

(c) Le théorème 24.10 de Banach-Steinhaus avec les semi-normes. Le fait qu'il s'applique bien à $\mathscr{D}(K)$.

(d) L'utilisation de cette version du théorème de Banach-Steinhaus dans la proposition 28.45. En particulier l'utilisation de la famille de fonctionnelles (28.119) et la preuve qu'elles sont continues.

(e) L'équivalence entre les deux topologies de la proposition 6.195 et le fait que cela s'applique bien à $\mathscr{D}(K)$.

(f) Et enfin l'utilisation de tout ça pour donner l'unique solution de l'équation de Schrödinger du théorème 30.41. Est-ce que l'énoncé de ce théorème est déjà correct ?

(g) En particulier la fonction (30.365) me semble être un petit bricolage.

(6) Peut-on permuter une application linéaire et continue avec une somme pas spécialement dénombrable ? En supposant que $\sum_{i\in I} f(v_i)$ existe, la proposition 9.187 semble dire que oui. Est-ce correct ?

Peut-on avoir un exemple de partie sommable $\{v_i\}_{i\in I}$ et d'application linéaire continue f telle que la partie $\{f(v_i)\}$ ne soit pas sommable ?

Peut-être ceci dans un espace de Hilbert. $v_i = \frac{1}{i^2}e_i$ puis $f(e_i) = i^3 e_i$.

(7) Soit une partie orthonormale *pas spécialement dénombrable* $\{u_i\}_{i\in I}$ d'un espace de Hilbert (pas spécialement séparable). Si

$$x = \sum_{i\in I} x_i u_i, \tag{1.28}$$

4. i.e. admet une base dénombrable de topologique, voir la proposition 6.48

puis-je prendre le produit scalaire avec u_k et le permuter avec la somme pour déduire que $x_k = \langle x, u_k \rangle$?
C'est ce que je fais dans la proposition 23.26.

(8) Théorème de point fixe et équation différentielle. Que penser de l'exemple 18.42 qui itère la contraction de Cauchy-Lipschitz pour résoudre $y'(t) = y(t)$, $y(0) = 1$? Est-ce que c'est générique comme comportement ? Est-ce que la convergence est efficace dans des cas moins triviaux ?

1.5.4 Question de numérique

(1) Différences finies. Il faut une analyse de consistance, stabilité et convergence du schéma à 9 points pour le laplacien donné par (33.146). Je crois qu'il est d'ordre 6, mais je n'en suis vraiment pas sûr.

1.5.5 Pour ma culture générale

Est-ce qu'il existe un domaine des mathématiques (logique formelle, théorie des catégories ?) qui ne fasse pas du tout intervenir la constante de Weiner (définition 41.1) ?

1.6 Comment contribuer et aider (LaTeX et programmation)

1.6.1 Niveau facile

Une simple recherche avec quelque mots clefs devait donner la réponse ; n'importe qui ayant tapé son mémoire de licence devrait pouvoir le faire.

(1) Comment faire en sorte que les mots commençant par «é» soient avec les «e» dans l'index, et non avant les «a» ?

1.6.2 Niveau moyen

Demande d'avoir un peu touché à du LaTeX pour personnaliser son document.

(1) L'environnement `example` est un sale hack pour placer le triangle. Faire mieux. Recopier l'environnement théorème pour lui changer son carré en triangle ?

(2) Dans le script `splittoc.py`, la procédure de création des morceaux de pdf peut être parallélisée. Et séparément, la procédure de recomposition des volumes peut également être parallélisée.

1.6.3 Niveau difficile

Je ne sais pas comment faire, et à mon avis il faudra innover.

(1) Si vous savez comment faire `pdf -> epub` pour créer un eBook, faites le moi savoir. Cahier des charges :
— libre, disponible sur Ubuntu
— en ligne de commande (en tout cas : exécutable depuis un script en python ou C++)
Attention : le Frido étant un truc assez compliqué, avant de répondre la première chose qui vous passe par la tête, assurez-vous que votre solution fait avancer les choses sur le Frido et non sur un petit document de test.
Nous en avons déjà un peu discuté sur `https://github.com/LaurentClaessens/mazhe/issues/13`. Il faudra entre autres faire un script qui remplace tous les environnements tizk des fichiers `*.pstricks` (désolé pour la convention de nommage historique) par un simple `includegraphics` du fichier `pdf` correspondant que l'on trouvera dans le répertoire `auto/pictures_tikz`.

(2) Écrire un script (en python ou autre) qui prend en argument deux numéros ou noms de chapitres et qui retourne l'ensemble des lignes du premier qui contient des `ref` ou `eqref` dont le label correspondant est dans le second.

Attention : il faut tenir compte de `input` de façon récursive.

Bonus : calculer le hash sha1 de chaque ligne du résultat et ne pas l'afficher si il se trouve dans la liste du fichier `commons.py`.

1.7 Sage est là pour vous aider

Il existe de nombreux logiciels de mathématique. Notre préféré est Sage pour une raison très précise : en tant que langage de programmation, Sage est python qui est un langage généraliste. La syntaxe et la structure de Sage ne sont pas *ad hoc* pour faire de math, et ce qu'on apprend en Sage peut être recyclé pour faire n'importe quoi : navigateur web, script de manipulation de texte, traitement d'image, réseau neuronaux, ...

Sage est un logiciel disponible pour l'épreuve de modélisation de l'agrégation de mathématique ; il y a donc de bonnes chances que vous en ayez l'usage.

1.7.1 Lancez-vous dans Sage

(1) Aller sur `http://www.sagenmath.org`,

(2) créer un compte,

(3) créer des feuilles de calcul et s'amuser !!

Il y a beaucoup de documentation sur le site officiel [5], et nous vous conseillons particulièrement le livre [5].

Si vous comptez utiliser régulièrement ce logiciel, je vous recommande *chaudement* de l'installer sur votre ordinateur.

1.7.2 Exemples de ce que Sage peut faire pour vous

Ce livre est émaillé de petits bouts de code en Sage montrant ses différentes fonctionnalités là où nous en avons besoin [6]. Voici une liste (non exhaustive) de ce que Sage peut faire pour vous.

(1) Calculer des limites de fonctions, exemples 40.1 et 40.2.

(2) Tracer des graphes de fonctions, exemple 40.2.

(3) Tracer des courbes en trois dimensions, voir exemple 12.144. Notez que pour cela vous devez installer aussi le logiciel Jmol. Pour Ubuntu, c'est dans le paquet `icedtea6-plugin`.

(4) Calculer des dérivées partielles de fonctions à plusieurs variables, voir exemple 40.3.

(5) Résoudre des systèmes d'équations linéaires. Voir les exemples 40.4 et 40.5. Lire aussi la documentation.

(6) Tout savoir d'une forme quadratique, voir exemple 40.6.

(7) Calculer la matrice Hessienne de fonctions de deux variables, déterminer les points critiques, déterminer le genre de la matrice Hessienne aux points critiques et écrire extrema de la fonctions (sous réserve d'être capable de résoudre certaines équations), voir les exemples 40.7 et 40.8.

(8) Indiquer une infinité de solutions à une équation en utilisant des paramètres, voir l'exemple 40.9. Pour les fonctions trigonométriques,

```
sage: solve(sin(x)/cos(x)==1,x,to_poly_solve=True)
[x == 1/4*pi + pi*z1]
sage: solve(sin(x)**2==cos(x)**2,x,to_poly_solve=True)
```

5. http://www.sagemath.org

6. Soit un vrai besoin comme tracer un graphique en 3D, soit de la paresse comme calculer une grosse dérivée.

```
[sin(x) == cos(x), x == -1/4*pi + 2*pi*z86, x == 3/4*pi + 2*pi*z84]
```

Notez l'option `to_poly_solve=true` dans `solve`.

(9) Calculer des dérivées symboliquement, voir exemple 40.10.

(10) Calculer des approximations numériques comme dans l'exemple 40.11.

(11) Calculer dans un corps de polynômes modulo comme $\mathbb{F}_p[X]/P$ où P est un polynôme à coefficients dans $\mathbb{F}_p$. Voir l'exemple 5.162.

Sage peut en général faire tout ce que vous êtes capable de faire à l'entrée en master et probablement bien plus, à la notable exception des limites à deux variables.

Remarque 1.31.
Sage peut toutefois vous induire en erreur si vous n'y prenez pas garde parce qu'il sait des choses en mathématique que vous ne savez pas. Par conséquent il peut parfois vous donner des réponses (mathématiquement exactes) auxquelles vous ne vous attendez pas. Voir par exemple 15.86 pour le logarithme de nombres négatifs. Et aussi ceci :

```
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: limit(1/x,x=0)
7 Infinity
8 sage: limit(1/x**2,x=0)
9 +Infinity
```

tex/sage/sageSnip017.sage

Sage fait une différence entre `Infinity` et `+Infinity` et donne

$$\lim_{x\to 0} \frac{1}{x} = \infty \tag{1.29}$$

ainsi que

$$\lim_{x\to 0} \frac{1}{x^2} = +\infty. \tag{1.30}$$

Chapitre 2

Construction des ensembles de nombres

2.1 Quelques éléments sur les ensembles

Nous supposons, comme point de départ de notre mathématique, une théorie naïve des ensemble comme acceptée. Sachez toutefois pour votre culture générale que la construction de la théorie des ensembles n'a absolument rien ni d'évident ni de naïf. Des paradoxes arrivent très facilement si la théorie des ensembles n'est pas bien fondée. Par exemple, que dire de l'ensemble des ensembles qui ne se contiennent pas ? Ou encore ceci. Soit l'ensemble des nombres naturels qui peuvent être décrits par un texte en français de 28 mots. C'est un ensemble fini. Prenez le maximum, ajoutez un. Boum. Nous avons écrit un texte de 28 mots qui décrit un nombre qui n'est pas dans l'ensemble des nombres descriptibles en 28 mots.

Bref. Nous supposons avoir une théorie des ensembles qui tient la route.

Définition 2.1.
Un ensemble est ***infini*** *s'il peut être mis en bijection avec un de ses sous-ensembles propres.*

2.1.1 Ensemble ordonné

2.2.
L'**axiome du choix** que nous acceptons peut s'énoncer comme ceci[6] : Étant donné un ensemble X d'ensembles non vides, il existe une fonction définie sur X, appelée fonction de choix, qui à chacun d'entre eux associe un de ses éléments.

Définition 2.3.
Une ***relation d'ordre*** *sur un ensemble E est une relation binaire (notée $\leqslant$) sur E telle que pour tous $x, y, z \in E$,*

réflexivité *: $x \leqslant x$*

antisymétrie *: $x \leqslant y$ et $y \leqslant x$ implique $x = y$*

transitivité *: $x \leqslant y$ et $y \leqslant z$ implique $x \leqslant z$.*

Définition 2.4.
Un ensemble ordonné est ***totalement ordonné*** *si deux éléments sont toujours comparables : si $x, y \in E$ alors nous avons soit $x \leqslant y$ soit $y \leqslant x$. Si les éléments ne sont pas tous comparables, nous disons que l'ordre est* ***partiel****.*

Définition 2.5.
Soit un ensemble ordonné $(E, \leqslant)$ et une partie A de E. Nous disons que $m \in A$ est un ***minimum*** *si pour tout $y \in E$, l'élément m est comparable à m et $m \leqslant y$.*

Définition 2.6.
Un ensemble ordonné est ***bien ordonné*** *si toute partie non vide possède un plus petit élément : si A est une partie de E, alors $\exists x \in A, \forall y \in A, x \leqslant y$.*

2.7.
Quelques remarques.

(1) L'inégalité stricte (définie par : $x < y$ si et seulement si $x \leqslant y$ et $x \neq y$) n'est pas une relation d'ordre parce qu'elle n'est pas réflexive.
(2) Nous verrons dans la remarque 2.53 que l'intervalle $[-1, 1]$ dans $\mathbb{R}$ n'est pas bien ordonné.
(3) Un ensemble bien ordonné est forcément totalement ordonné parce que toutes les parties de la forme $\{x, y\}$ possèdent un minimum. Par conséquent x et y doivent être comparables : $x \leqslant y$ ou $y \leqslant x$.

La notion d'ordre permet d'introduire la notion d'intervalle.

Définition 2.8.
Soit un ensemble totalement ordonné $(E, \leqslant)$. Un ***intervalle*** *de E est une partie telle que tout élément compris entre deux éléments de la partie soit dedans. En formule, la partie I de E est un intervalle si*

$$\forall a, b \in I, (a \leqslant x \leqslant b) \Rightarrow x \in I.$$

2.1.2 Lemme de Zorn

Définition 2.9 (Ensemble inductif[7])**.**
Un ensemble est ***inductif*** *si tout sous-ensemble totalement ordonné admet un majorant.*

Si E est un ensemble, l'inclusion est un ordre sur l'ensemble des partie de E, mais pas un ordre total parce que si X, Y sont des parties de E, alors nous n'avons pas automatiquement soit $X \subset Y$ ou $Y \subset X$.

Lemme 2.10 (Lemme de Zorn)**.**
Tout ensemble ordonné inductif non vide admet un maximum.

Le point intéressant de ce lemme est que le majorant soit un maximum, c'est à dire qu'il appartienne à l'ensemble.

Proposition 2.11.
Si S est un ensemble infini alors il existe une bijection $\varphi\colon \{1, 2\} \times S \to S$.

2.1.3 Complémentaire

Soit E, un ensemble et A, une partie de E (c'est à dire un sous-ensemble de E). Nous désignons par $\complement A$ désigne le **complémentaire** de l'ensemble A dans E. Il s'agit de l'ensemble des points de E qui ne font pas partie de A :

$$\complement A = E \backslash A = \{x \in E \text{ tel que } x \notin A\}. \tag{2.1}$$

Nous allons aussi régulièrement noter le complémentaire de A par A^c.

Lemme 2.12.
Quelques propriétés à propos des complémentaires. Si E est un ensemble et si A et B sont des sous-ensembles de E, nous avons

(1) $\complement\complement A = A$, *en d'autres termes,* $E \backslash (E \backslash A) = A$,
(2) $\complement(A \cap B) = \complement A \cup \complement B$,
(3) $\complement(A \cup B) = \complement A \cap \complement B$,
(4) $A \backslash B = A \cap \complement B$.

Définition 2.13 (différence symétrique)**.**
Si A et B sont des ensembles, l'ensemble $A\Delta B$ est la ***différence symétrique*** *d'ensembles :*

$$A\Delta B = (A \cup B)\backslash(A \cap B). \tag{2.2}$$

C'est l'ensemble des éléments étant soit dans A soit dans B mais pas dans les deux.

Lemme 2.14.
Si A et B sont des ensembles nous avons

(1) $A^c\Delta B^c = A\Delta B$.

(2) $(A\Delta B)\Delta B = A$.

Démonstration. La première assertion provient de l'égalité $X^c\backslash Y^c = Y\backslash X$ (qu'on montre de façon classique, éventuellement en séparant les cas suivant que B est inclus à A ou non) :

$$A^c\Delta B^c = (A^c \cup B^c)\backslash(A^c \cap B^c) = (A \cap B)^c\backslash(A \cup B)^c = (A \cup B)\backslash(A \cap B) = A\Delta B. \tag{2.3}$$

Pour la seconde assertion, il faut remarquer que $(A\Delta B) \cup B = A \cup B$ et que $(A\Delta B) \cap B = B\backslash A$, donc

$$(A\Delta B)\Delta B = (A \cup B)\backslash(B\backslash A) = A. \tag{2.4}$$

□

2.1.4 Relations d'équivalence

Définition 2.15.
Si E est un ensemble, une ***relation d'équivalence*** *sur E est une relation $\sim$ qui est à la fois*

réflexive *$x \sim x$ pour tout $x \in E$,*

symétrique *$x \sim y$ si et seulement si $y \sim x$;*

transitive *si $x \sim y$ et $y \sim z$, alors $x \sim z$.*

Par exemple, sur l'ensemble de tous les polygones du plan, la relation «a le même nombre de côtés» est une relation d'équivalence. Plus précisément, si P et Q sont deux polygones, nous disons que $P \sim Q$ si et seulement si P et Q ont le même nombre de côtés. Cela est une relation d'équivalence :

— un polygone P a toujours le même nombre de côtés que lui-même : $P \sim P$;
— si P a le même nombre de côtés que Q ($P \sim Q$), alors Q a le même nombre de côtés que P ($Q \sim P$) ;
— si P a le même nombre de côtés que Q ($P \sim Q$) et que Q a le même nombre de côtés que R ($Q \sim R$), alors P a le même nombre de côtés que R ($P \sim R$).

Soit f une application entre deux ensembles E et F. Nous définissons une relation d'équivalence sur E par

$$x \sim y \Leftrightarrow f(x) = f(y). \tag{2.5}$$

Nous notons par $\pi\colon E \to E/\sim$ la projection canonique. L'application

$$\begin{aligned} g\colon E/\sim &\to F \\ [x] &\mapsto f(x) \end{aligned} \tag{2.6}$$

est bien définie et injective. Elle n'est pas surjective tant que f ne l'est pas. La **décomposition canonique** de f est

$$f = g \circ \pi. \tag{2.7}$$

2.2 Les naturels

Toutes les constructions sont faites dans [8]. Les résultats énoncés ici sont utilisés plus bas et servent de guide à un éventuel contributeur qui voudrait écrire une partie sur la construction de $\mathbb{N}$ et $\mathbb{Z}$. Nous espérons que des preuves se trouvent dans [8]. En tout cas, le lecteur est invité à ne pas les prendre pour évidents.

Définition 2.16.
Un ensemble est ***dénombrable*** *s'il peut être mis en bijection avec* $\mathbb{N}$*. Il est* ***non dénombrable*** *s'il est infini et ne peut pas être mis en bijection avec* $\mathbb{N}$*.*

Une chose vraiment amusante avec cette définition que l'on met en rapport avec la définition 2.1, c'est qu'un ensemble fini n'est ni dénombrable ni non dénombrable.

Proposition 2.17.
Toute partie d'un ensemble fini est finie, et toute partie d'un ensemble dénombrable est finie ou dénombrable.

2.2.1 La construction

Cette section est vide, insuffisamment détaillée ou incomplète. Votre aide est la bienvenue ! Comment faire ?

2.2.2 Quelque autres résultats

Nous supposons ici connaître, et avoir démontré les rudiments du calcul dans $\mathbb{N}$. En particulier les produits remarquables, et les implications comme $a > b$ implique qu'il n'existe pas de x tel que $a + b = b$.

La proposition suivante donne une bijection explicite entre $\mathbb{N}$ et $\mathbb{N} \times \mathbb{N}$. Elle n'a rien de transcendante, mais je ne résiste pas à la donner ici parce qu'elle est utilisée dans l'article *Un peu de programmation transfinie* de David Madore[1]. Son utilité est de pouvoir créer un language de programmation pouvant traiter des pairs d'entiers rien qu'en traitant des entiers.

Proposition 2.18 (Une bijection $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$)**.**
La fonction

$$\begin{aligned} f\colon \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ (x, y) &\mapsto \begin{cases} y^2 + x & si\ x < y \\ x^2 + x + y & si\ y \leqslant x. \end{cases} \end{aligned} \tag{2.8}$$

est une bijection.

Démonstration. Il s'agit de prouver qu'elle est injective et surjective. Dans la suite, tous les nombres sont des entiers positifs. D'ailleurs nous n'avons encore rien défini d'autre.

f est injective Pour $k \in \mathbb{N}$ donné, nous allons prouver que

(1) l'équation $f(x, y) = k$ possède au maximum une solution avec $x < y$,

(2) l'équation $f(x, y) = k$ possède au maximum une solution avec $y \leqslant x$,

(3) si $k = y^2 + x$ avec $x < y$ alors il est impossible que $k = x'^2 + y'$ avec $y \leqslant x$.

Supposons $k = y^2 + x$ avec $x < y$. Alors aucun y' de la forme $y + s$ ne peut être solution parce que

$$y^2 + x = k = (y + s)^2 + x' = y^2 + 2sy + s^2 + x', \tag{2.9}$$

ce qui impliquerait $x' = x - 2sy - s^2 < 0$. Il ne peut donc pas y avoir deux solutions.

1. Et comme j'aime beaucoup cet article, il me fallait une excuse pour le placer ici.
`http://www.madore.org/~david/weblog/d.2017-08-18.2460.html`.

Supposons de la même manière que $k = x^2 + x + y$. Alors il n'existe pas de solutions avec $x' = x + s$ parce que cela donnerait

$$(x + s)^2 + (x + s) + y' = x^2 + x + y, \tag{2.10}$$

et donc

$$y' = y - 2sx - s^2 - s < 0. \tag{2.11}$$

f **est surjective** Soit $a \in \mathbb{N}$ nous allons montrer que tous les nombres de a^2 à $(a+1)^2$ sont des images de f. D'abord $a^2 = f(0,)$, ensuite les nombres

$$f(1, a), f(2, a), \ldots, f(a - 1, a) \tag{2.12}$$

prennent les valeurs $a^2+1, \ldots, a^2+a-1$. Après, $f(a, 0) = a^2+a$ et les nombres $f(a, 1), \ldots, f(a, a)$ prennent les valeurs de $a^2 + a + 1$ à $a^2 + 2a = (a + 1)^2 - 1$.

□

Sachez que cette fonction s'étend aux ordinaux (mais là ce n'est plus pour rigoler).

2.3 Les entiers

Lemme 2.19.
Toute partie bornée de $\mathbb{Z}$ possède un plus grand élément.

2.4 Quelques structures algébriques

Définition 2.20 (Anneau[9])**.**
*Un **anneau** est un triplet $(A, +, \cdot)$ avec les conditions*

(1) $(A, +)$ est un groupe abélien. Nous notons 0 le neutre.

(2) La multiplication est associative et nous notons 1 le neutre.

(3) La multiplication est distributive par rapport à l'addition.

Définition 2.21.
*Un **corps** est un anneau non réduit à $\{0\}$ dans lequel tout élément non nul est inversible.*

Remarque 2.22.
Un anneau est ce qu'on appelle «*ring*» en anglais. Un corps est en anglais «*field*». De plus le mot «*field*» comprend la commutativité. Donc certains utilisent le mot «corps» pour dire «corps commutatif» et parlent alors d'anneau *à division* pour parler de corps non commutatifs.

Définition 2.23 (Morphisme d'anneaux)**.**
*Si $(A, +, \cdot)$ et $(B, +, \cdot)$ sont des anneaux, un **morphisme** est une application $f : A \to B$ telle que*

(1) $f(a + b) = f(a) + f(b)$

(2) $f(a \cdot b) = f(a) \cdot f(b)$

(3) $f(1) = 1$.

Étant bien entendu que les significations de 1, $+$ et $\cdot$ sont différentes à gauche et à droite.

*Un **isomorphisme** est un morphisme bijectif.*

*Un corps étant un anneau sans plus de structure, un **morphisme de corps** n'est qu'un morphisme des anneaux.*

Définition 2.24 (Idéal dans un anneau)**.**
*Un sous ensemble $I \subset A$ est un **idéal à gauche** à gauche si*

(1) I est un sous-groupe pour l'addition,

(2) pour tout $a \in A$, $aI \subset I$.

Lorsqu'un ensemble est idéal à gauche et à droite, nous disons que c'est un ***idéal bilatère****. Lorsque nous parlons d'idéal sans précisions, nous parlons d'idéal bilatère.*

Définition 2.25.
Ordre et choses reliées dans un corps.

(1) Un corps $\mathbb{K}$ est ***totalement ordonné*** *s'il existe une relation d'ordre total*[2] *tel que*

(a) $x \leqslant y$ implique $x + z \leqslant y + z$ pour tout $x, y, z \in \mathbb{K}$

(b) $x \geqslant 0$ et $y \geqslant 0$ implique $xy \geqslant 0$.

(2) Si $\mathbb{K}$ est un corps totalement ordonné, nous y définissons la valeur absolue par

$$|x| = \begin{cases} x & si\ x \geqslant 0 \\ -x & si\ x \leqslant 0. \end{cases} \tag{2.13}$$

(3) La suite (x_n) dans le corps totalement ordonné $\mathbb{K}$ est ***de Cauchy*** *si pour tout $\epsilon \in \mathbb{K}^+$, il existe $N \in \mathbb{N}$ tel que si $p, q \geqslant N$ alors $|x_p - x_q| \leqslant \epsilon$.*

(4) La suite (x_n) dans le corps totalement ordonné $\mathbb{K}$ est ***convergente*** *s'il existe $q \in \mathbb{K}$ tel que pour tout $\epsilon \in \mathbb{K}^+$, il existe N tel que si $k \geqslant N$ alors $|x_k - q| \leqslant \epsilon$.*

(5) Un corps $\mathbb{K}$ est ***archimédien*** *s'il est totalement ordonné et si pour tout $x, y \in \mathbb{K}$ avec $x > 0$, il existe $n \in \mathbb{N}$ tel que $nx \geqslant y$.*

(6) Un corps totalement ordonné est ***complet*** *si toute suite de Cauchy dans $\mathbb{K}$ y est convergente.*

Remarque 2.26.
Nous étudierons plus tard la notion de caractéristique d'un anneau[3] ou d'un corps. Tout anneau totalement ordonné est nécessairement de caractéristique nulle, lemme 4.22.

Parmi ces définitions, celles de suite convergente, de Cauchy et de corps complet seront utilisées dans le cas de $\mathbb{Q}$ (et de $\mathbb{R}$ pour la complétude). Elles seront prouvées être équivalentes aux définitions topologiques dans le cas particulier de $\mathbb{R}$ et $\mathbb{Q}$ lorsque la topologie métrique sera définie. Dans cet état d'esprit nous n'allons pas démontrer tout de suite que $\mathbb{R}$ est un corps complet. Nous allons directement démontrer que c'est un espace topologique complet.

Lemme 2.27 (Propriétés de la valeur absolue)**.**
Soit $\mathbb{K}$ un corps totalement ordonné. Si $x, y \in \mathbb{K}$ alors

(1) Si $x \geqslant 0$ alors $-x \leqslant 0$.

(2) $|x| \geqslant 0$

(3) $|x| = 0$ si et seulement si $x = 0$

(4) $|x + y| \leqslant |x| + |y|$.

Démonstration. Point par point

(1) Nous partons de $x \geqslant 0$ et nous ajoutons $-x$ des deux côtés en profitant de la définition d'un corps totalement ordonné : $x - x \geqslant -x$ et donc $0 \geqslant -x$, c'est à dire $-x \leqslant 0$.

(2) Si $x \geqslant 0$ alors c'est vrai. Sinon, $x \leqslant 0$ et $|x| = -x \geqslant 0$ par le point (1).

(3) Si $x = 0$ alors $x = -x$ et $|x| = 0$. Au contraire si $x \neq 0$ alors $-x \neq 0$ et que x soit positif ou négatif, nous aurons toujours $\pm x \neq 0$.

(4) Nous supposons que $x \leqslant y$ et nous distinguons divers cas suivant la positivité de x et y.

(a) Si $x, y \geqslant 0$. Dans ce cas, $x + y \geqslant y \geqslant 0$, donc $|x + y| = x + y = |x| + |y|$.

(b) Si $x, y \leqslant 0$. Dans ce cas, $x + y \leqslant 0$ et nous avons $|x + y| = -x - y = |x| + |y|$.

2. Définition 2.4.
3. définition 4.18

(c) Si $x \leqslant 0$ et $y \geqslant 0$. Nous subdivisons encore en deux cas suivant que $x+y$ est positif ou négatif. Si $x+y \geqslant 0$, alors nous écrivons successivement

$$x \leqslant 0 \tag{2.14a}$$
$$x+y \leqslant y \leqslant y+|x| = |x|+|y| \tag{2.14b}$$

et donc $|x+y| = x+y \leqslant |x|+|y|$.
Nous supposons à présent que $x \leqslant 0$, $y \geqslant 0$ et $x+y \leqslant 0$. Dans ce cas il suffit d'écrire $|x+y| = |(-x)+(-y)|$ pour retomber dans le cas précédent à inversion près de x et y.

□

Remarque 2.28.
La partie (4) est très importante parce que c'est elle qui fera presque toutes les majorations dont nous aurons besoin en analyse. En effet elle donne l'inégalité triangulaire de la façon suivante : si $x, y, z \in \mathbb{K}$ nous avons

$$|x-y| = |(x-z)+(z-y)| \leqslant |x-z|+|z-y|. \tag{2.15}$$

2.5 Les rationnels

Toutes les constructions sont faites dans [8]. Les résultats énoncés ici sont utilisés plus bas et servent de guide à un éventuel contributeur qui voudrait écrire une partie sur la construction de $\mathbb{Q}$ à partir de $\mathbb{Z}$. Nous espérons que des preuves se trouvent dans [8]. En tout cas, le lecteur est invité à ne pas les prendre pour évidents.

Lemme 2.29.
Tout rationnel est majoré par un naturel.

La proposition 4.59 dit que $\mathbb{Q}$ est le corps des fractions de l'anneau $\mathbb{Z}$.

2.5.1 Suites de Cauchy dans les rationnels

Proposition 2.30 ([8])**.**
Principales propriétés des suites de Cauchy dans $\mathbb{Q}$.

(1) Toute suite convergente est de Cauchy[4]*.*

(2) Toute suite de Cauchy est bornée.

(3) Si $x_n \to 0$ et si (y_n) est bornée, alors $x_n y_n \to 0$

(4) Si (x_n) et (y_n) sont de Cauchy alors (x_n+y_n), (x_n-y_n) et $(x_n y_n)$ sont également de Cauchy.

(5) Si $x_n \to a$ et $y_n \to b$ alors $x_n+y_n \to a+b$, $x_n-y_n \to a-b$ et $x_n y_n \to ab$.

(6) Soit (x_n) une suite de Cauchy qui ne converge pas vers zéro. Alors il existe n_0 tel que la suite $\left(\frac{1}{x_n}\right)_{n\geqslant n_0}$ soit de Cauchy.

Démonstration. Point par point.

(1) Soit (x_n) une suite dans $\mathbb{Q}$ qui converge vers $x \in \mathbb{Q}$. Soit aussi[5] $\epsilon > 0$. Il existe $N_\epsilon \in \mathbb{N}$ tel que $n > N_\epsilon$ implique $|x_n - x| \leqslant \epsilon$. Si $p, q \geqslant N_{\epsilon/2}$ alors

$$|x_p - x_q| \leqslant |x_p - x| + |x - x_q| \leqslant \epsilon. \tag{2.16}$$

Donc la suite (x_n) est de Cauchy.

4. Et non la réciproque, qui sera justement la grande innovation des nombres réels.
5. Le ϵ que nous prenons maintenant est dans $\mathbb{Q}$, et ce sera toujours ainsi dans les prochaines pages ; nous ne le répéterons pas à chaque fois.

(2) Soit (x_n) une suite de Cauchy dans $\mathbb{Q}$. Avec $\epsilon = 1$ dans la définition, si $q > N_1$, nous avons

$$|x_q - x_{N_1}| \leqslant 1. \tag{2.17}$$

Et donc pour tout q plus grand que N_1, $x_N - 1 \leqslant x_q \leqslant x_N + 1$, ou encore, pour tout n :

$$|x_n| \leqslant \max\{|x_1|, |x_2|, \ldots, |x_N|, |x_N + 1|\}. \tag{2.18}$$

La suite est donc bornée.

(3) Soit $\epsilon > 0$. Les hypothèses disent qu'il existe un N tel que $|x_n| \leqslant \epsilon$ dès que $n \geqslant N$. Et il existe aussi $M \geqslant 0$ tel que $|y_n| \leqslant M$ pour tout n. Du coup, lorsque $n \geqslant N$ nous avons $|x_n y_n| \leqslant M\epsilon$.

(4) En ce qui concerne la somme,

$$|x_p + y_p - x_q - y_q| \leqslant |x_p - x_q| + |y_p - y_q|. \tag{2.19}$$

Soit N_1 tel que si $p, q \geqslant N_1$ alors $|x_p - x_q| \leqslant \epsilon$ et N_2 de même pour la suite (y_n). En prenant $N = \max\{N_1, N_2\}$, la somme (2.19) est plus petite que 2ϵ dès que $p, q \geqslant N$.
Passons à la démonstration du fait que le produit de deux suites de Cauchy est de Cauchy. Les suites (x_n) et (y_n) sont bornées et quitte à prendre le maximum, nous disons qu'elles sont toutes les deux bornées par le nombre M : pour tout n nous avons $|x_n| \leqslant M$ et $|y_n| \leqslant M$. Nous avons :

$$|x_p y_p - x_q y_q| \leqslant |x_p y_p - x_q y_p| + |x_q y_p - x_q y_q| \leqslant |y_p||x_p - x_q| + |x_q||y_p - y_q|. \tag{2.20}$$

Vu que (x_n) et (y_n) sont de Cauchy, si p et q sont assez grands, les deux différences sont majorées par ϵ et nous avons

$$|x_p y_p - x_q y_q| \leqslant M\epsilon + M\epsilon = 2M\epsilon, \tag{2.21}$$

ce qui prouve que $(x_n y_n)$ est de Cauchy.

(5) En ce qui concerne la somme, nous pouvons tout de suite calculer

$$|x_n + y_n - (a + b)| \leqslant |x_n - a| + |y_n - b|. \tag{2.22}$$

Il existe une valeur de n à partir de laquelle le premier terme est plus petit que ϵ et une à partir de laquelle le second terme est plus petit que ϵ. En prenant le maximum des deux, la somme est plus petite que 2ϵ.
En ce qui concerne le produit,

$$|x_n y_n - ab| \leqslant |x_n y_n - a y_n| + |a y_n - ab| \leqslant |y_n||x_n - a| + |a||y_n - b|. \tag{2.23}$$

Les suites $|x_n - a|$ et $|y_n - b|$ convergent vers zéro ; la suite (y_n) est bornée parce que convergente (combinaison des points (1) et (2)) et a (la suite constante) est également bornée. Donc par le point (3), nous avons

$$y_n|x_n - a| + a|y_n - b| \to 0. \tag{2.24}$$

Au passage nous avons également utilisé la propriété de la somme que nous venons de démontrer.

(6) Soit (x_n) une suite de Cauchy dans $\mathbb{Q}$ ne convergeant pas vers zéro : il existe $\alpha > 0$ tel que pour tout $N \in \mathbb{N}$, il existe $n \geqslant N$ tel que $|x_n| > \alpha$. Mais notre suite est de Cauchy, donc il existe $n_0 \in \mathbb{N}$ tel que si $p, q \geqslant n_0$ alors

$$|x_p - x_q| \leqslant \frac{\alpha}{2}. \tag{2.25}$$

En fixant $N = n_0$, on obtient un naturel $n \geqslant n_0$ tel que $|x_n| \geqslant \alpha$. De plus, comme la suite est de Cauchy, si $p > n$ nous avons aussi $|x_n - x_p| \leqslant \frac{\alpha}{2}$. Cela implique $|x_p| \geqslant \frac{\alpha}{2}$ et en particulier $x_p \neq 0$.
Nous venons de prouver que la suite ne s'annule plus à partir de l'indice n, et même que $|x_k| \geqslant \alpha/2$ pour $x \geqslant n$. La suite $(1/x_k)_{k\geqslant n}$ est donc bien définie.
Soit $\epsilon > 0$. Soit n_0 tel que $|x_p - x_q| < \epsilon$ pour tout $p, q > n_0$. Soit K plus grand que n_0 et que n. En prenant $p, q \geqslant K$, nous avons $|x_p| > \frac{\alpha}{2}$ et $|x_q| > \frac{\alpha}{2}$. Nous en déduisons que

$$\left|\frac{1}{x_p} - \frac{1}{x_q}\right| \leqslant \frac{|x_q - x_p|}{|x_p x_q|} \leqslant \frac{4}{\alpha^2}|x_q - x_p| \leqslant \frac{4}{\alpha^2}\epsilon. \tag{2.26}$$

Donc $\left(\frac{1}{x_n}\right)$ est de Cauchy.

□

2.5.2 Insuffisance des rationnels

Nous allons voir qu'il n'existe pas de nombres rationnels x tels que $x^2 = 2$, mais que pourtant il existe une infinité de suites de rationnels (x_n) tels que $x_n^2 \to 2$.

Lemme 2.31.
Un entier x est pair si et seulement si l'entier x^2 est pair.

Démonstration. Si x est un nombre pair, alors il existe un entier a tel que $x = 2a$ alors $x^2 = 4a^2$ est pair.

Inversement, si x est impair alors il existe un entier a tel que $x = 2a + 1$ et alors $x^2 = 4a^2 + 4a + 1 = 2(2a^2 + 2a) + 1$ est impair. □

Proposition 2.32 (Irrationalité de $\sqrt{2}$)**.**
Il n'existe pas de fractions d'entiers dont le carré soit égal à 2.

Le théorème 12.24 nous dira que tous les $\sqrt{n}$ sont irrationnels dès que n n'est pas un carré parfait.

Démonstration. Nous supposons que la fraction d'entiers a/b est telle que $a^2/b^2 = 2$, et nous allons construire une suite d'entiers strictement décroissante et strictement positive, ce qui est impossible.

Grâce au lemme 2.31 nous avons successivement les affirmations suivantes :
— $\frac{a^2}{b^2} = 2$ avec $a \neq 0$ et $b \neq 0$.
— $a^2 = 2b^2$, donc a^2 est pair.
— a est alors pair et a^2 est divisible par 4. Soit $a^2 = 4k$.
— $4k/b^2 = 2$, donc $4k = 2b^2$, donc $b^2 = 2k$ et b^2 est pair.
— Nous déduisons que b est pair.

La fraction $\frac{a/2}{b/2}$ est alors une nouvelle fraction d'entiers dont le carré vaut 2. En procédant de la même façon, en remplaçant a par $a/2$ et b par $b/2$, on obtient que la fraction d'entiers $\frac{a/4}{b/4}$ a la même propriété.

En particulier, tous les nombres de la forme $a/2^n$ sont des entiers. Ils forment une suite strictement décroissante d'entiers strictement positifs. Impossible, me diriez vous ? Et vous auriez bien raison : toute partie non vide de $\mathbb{N}$ admet un plus petit élément [6]. Il n'y a donc pas de fractions d'entiers dont le carré vaut 2. □

Proposition 2.33 ([1])**.**
Soit la suite de rationnels (x_n) définie par $x_0 \in \mathbb{Q}^+$ et

$$x_{n+1} = x_n + \frac{x_n^2 - 2}{2x_n}. \tag{2.27}$$

Alors :

6. Voir [8], et attention : ce n'est pas tout à fait évident.

(1) En posant $y_n = x_n^2$ nous avons $y_n \to 2$.

(2) La suite (x_n) est de Cauchy.

(3) La suite (x_n) ne converge pas dans $\mathbb{Q}$.

Démonstration. Tout d'abord un petit calcul montre que

$$x_{n+1}^2 = 2 + \frac{(y_n - 2)^2}{4y_n}, \tag{2.28}$$

c'est à dire que quelle que soit la valeur de x_0, dès le x_1, les valeurs de y_n sont plus grandes que 2. Notons qu'il n'est pas possible d'avoir $y_n = 2$. Nous pouvons donc supposer $y_n > 2$ pour tout n. Alors en posant $y_n = 2 + s$ nous avons

$$y_{n+1} = 2 + \frac{s^2}{8 + 4s}. \tag{2.29}$$

Si $s < 1$ alors $\frac{s^2}{8+4s} < s^2$ et en réalité le processus $s \mapsto s^2/(8 + 4s)$ tend très vite vers zéro. Nous devons donc montrer qu'il existe un n tel que $y_n = 2 + s$ avec $s < 1$.

Montrons pour cela que si $n < s < 2n$ alors[7]

$$\frac{s^2}{8 + 4s} < n \tag{2.30}$$

En effet si $n < s < 2n$ nous pouvons majorer le numérateur par $4n^2$ et minorer le dénominateur par $4n$.

Prouvons à présent le résultat. Pour un $n > 1$ nous avons

$$y_{n+1} = 2 + \frac{s}{8 + 4s} \tag{2.31}$$

avec $s > 0$. Nous considérons $n_0 \in \mathbb{N}$ tel que $n_0 < s < 2n_0$. Alors

$$y_{n+2} = 2 + s_1 \tag{2.32}$$

avec $s_1 < n$. Il existe donc n_1 tel que $n_1 < s_1 < 2n_1$ et $n_1 < n_0$. Nous avons alors $y_{n+3} = 2 + s_2$ avec $s_2 < n_1 < n_0$.

Nous construisons ainsi des suites (s_i) et (n_i) telles que

$$y_{n+k} = 2 + s_k \tag{2.33}$$

avec $s_k < n_{k-1} < n_{k-2} < \cdots < n_0$. En procédant ainsi au maximum $n - 1$ fois nous avons $s_k < 1$. À partir du moment où $y_{n+k} = 2 + s_k$ avec $s_k < 1$, nous avons déjà vu qu'il est certain que $y_n \to 2$.

Prouvons que la suite (x_n) est de Cauchy. Vu que $x_n^2 \to 2$ nous avons $x_n \geqslant 1$ à partir d'une certaine valeur de n. Soit $\epsilon > 0$; comme (x_n^2) converge, par la proposition 2.30(1), elle est de Cauchy, et il existe n_0 tel que $|x_p^2 - x_q^2| < \epsilon$ pour $p, q \geqslant n_0$. Alors, pour p, q plus grands que n_0 et n, nous avons

$$|x_p - x_q| = \frac{|x_p^2 - x_q^2|}{|x_p + x_q|} \leqslant \frac{\epsilon}{2}. \tag{2.34}$$

Cela prouve que la suite (x_n) est de Cauchy dans $\mathbb{Q}$.

Enfin supposons que $x_n \to a \in \mathbb{Q}$. Dans ce cas nous aurions $x_n^2 \to a^2 = 2$ (proposition 2.30(5)). Mais nous savons par la proposition 2.32 que $a^2 = 2$ est impossible dans $\mathbb{Q}$. □

Bref, la suite (x_n) est de Cauchy, ne converge pas et a son carré qui converge vers 2.

7. Voir la remarque 2.34 pour comprendre d'où vient l'idée de cette majoration.

Remarque 2.34.
Vu que nous n'avons pas encore défini les réels, ce qui suit n'est que informel. Ce qui a motivé la majoration (2.30) est la résolution de l'inéquation

$$\frac{s^2}{8+4s} < n \tag{2.35}$$

qui donne les bornes $2n \pm \sqrt{n^2+2n}$ dont une est toujours négative, et l'autre plus grande que $2n$.

Vous en voulez encore une ?

Lemme 2.35 (Série géométrique, voir aussi l'exemple 10.62)**.**
Si $q \in \mathbb{Q}$ et $p \in \mathbb{N}$ nous avons

$$\sum_{k=0}^{p} q^k = \frac{1-q^{p+1}}{1-q}. \tag{2.36}$$

Démonstration. En posant $S_p = 1 + q + q^2 + \cdots + q^p$, nous avons $S_p - qS_p = 1 - q^{p+1}$ et donc

$$S_p = \sum_{k=0}^{p} q^k = \frac{1-q^{p+1}}{1-q}. \tag{2.37}$$

□

Proposition 2.36.
La suite donnée par

$$x_n = 1 + \frac{1}{1!} + \cdots + \frac{1}{n!} \tag{2.38}$$

est de Cauchy et ne converge pas dans $\mathbb{Q}$.

Démonstration. Si $p > q > 0$ nous avons

$$\begin{aligned}
x_p - x_q &= \sum_{k=q+1}^{p} \frac{1}{k!} && \text{(2.39a)}\\
&\leqslant \sum_{k=q+1}^{p} \frac{1}{(q+1)!}\frac{1}{(q+1)^{k-q-1}} && \text{(2.39b)}\\
&\leqslant \frac{1}{(q+1)!} \lim_{p\to\infty} \sum_{k=0}^{p} \frac{1}{(q+1)^k} && \text{(2.39c)}\\
&= \frac{1}{(q+1)!}\frac{1}{1-\frac{1}{q+1}} && \text{(2.39d)}\\
&= \frac{1}{(q+1)!}\frac{q+1}{q} && \text{(2.39e)}\\
&= \frac{1}{q!q}. && \text{(2.39f)}
\end{aligned}$$

Justifications :

— Pour (2.39b), il s'agit de remplacer dans $k!$ tous les facteurs plus grands que $(q+1)$ par $q+1$. Cela rend le dénominateur plus petit.

— Pour (2.39c), il y a une inégalité parce que la suite $p \mapsto \sum_{k=0}^{p} 1/(q+1)^k$ est une suite strictement croissante.

— Pour (2.39d), le lemme 2.35 donne la valeur de la somme finie. En ce qui concerne la limite, nous avons demandé $p > q > 0$ et donc $q + 1 > 1$. Dans ce cas la limite fonctionne.

Cette inégalité une fois établie nous permet de prouver les assertions. La suite (x_n) est de Cauchy car, pour tout $\epsilon \in \mathbb{Q}$ s'écrivant $\epsilon = \frac{a}{b}$ avec $a, b \in \mathbb{N}$, en prenant $p, q > b$, nous avons

$$x_p - x_q \leqslant \frac{1}{b!b} < \frac{1}{b} < \frac{a}{b} = \epsilon. \tag{2.40}$$

Montrons par l'absurde que cette suite ne converge pas dans $\mathbb{Q}$. Supposons donc que $\lim_{n\to\infty} x_n = \frac{a}{b} \in \mathbb{Q}$. Pour tout $p > q$ nous avons établi

$$0 < x_p - x_q < \frac{1}{qq!}. \tag{2.41}$$

Prenons la limite $p \to \infty$; par stricte croissance de la suite, les inégalités restent strictes :

$$0 < \frac{a}{b} - x_q < \frac{1}{qq!}. \tag{2.42}$$

Si $n > b$ alors nous pouvons écrire

$$\frac{a}{b} - x_n = \frac{\alpha}{n!} \tag{2.43}$$

avec $\alpha \in \mathbb{Z}$ parce que le dénominateur commun entre $\frac{a}{b}$ et x_n est dans $n!$. En prenant donc $q > n$ dans (2.42) nous pouvons écrire

$$0 < \frac{\alpha}{q!} < \frac{1}{qq!}, \tag{2.44}$$

c'est à dire $0 < \alpha < \frac{1}{q}$, ce qui est impossible pour $\alpha \in \mathbb{Z}$. □

2.6 Les réels

Une construction des réels via les coupures de Dedekind est donnée dans [10].

2.37.
La construction des réels va nécessiter un petit «bootstrap» au niveau de la topologie. En effet la notion de suite de Cauchy est une notion topologique (définition 6.37) alors que la topologie métrique (celle entre autres de $\mathbb{Q}$) ne sera définie que par le théorème 6.19. Nous avons donc dû définir en la définition 2.25 *ex nihilo* les notions de

— suite de Cauchy
— suite convergente
— complétude

Nous allons ensuite construire $\mathbb{R}$ comme ensemble de classes d'équivalence de suites de Cauchy dans $\mathbb{Q}$. Ce ne sera que plus tard, après avoir défini la topologie métrique que nous allons voir que sur $\mathbb{R}$, ces trois notions coïncident avec celles topologiques [8]. Et par conséquent que $\mathbb{R}$ sera un espace topologique complet [9].

Dans cette optique, il est intéressant de lire ce que dit Wikipédia à propos des suites de Cauchy dans l'article consacré à la construction des nombres réels.

2.6.1 L'ensemble

Soit $\mathcal{E}$ l'ensemble des suites de Cauchy [10] dans $\mathbb{Q}$. Soit aussi l'ensemble $\mathcal{E}_0$ constituée des suites qui convergent vers zéro [11].

En posant

$$x + y = (x_n + y_n) \tag{2.45}$$

et

$$xy = (x_n y_n), \tag{2.46}$$

l'ensemble $\mathcal{E}$ devient un anneau [12] commutatif dont le neutre de l'addition est la suite constante $x_n = 0$ et le neutre pour la multiplication est la suite constante $x_n = 1$.

8. Proposition 6.52.
9. Théorème 6.57.
10. Définition 2.25(3)
11. Nous rappelons qu'à ce niveau nous n'avons pas encore prouvé que toutes les suites de Cauchy convergent.
12. Définition 2.20.

Proposition 2.38.
La partie $\mathcal{E}_0$ est un idéal[13] *de l'anneau $\mathcal{E}$.*

Démonstration. Nous savons par la proposition 2.30(1) que les suites convergentes sont de Cauchy ; par conséquent $\mathcal{E}_0 \subset \mathcal{E}$.

L'ensemble structuré $(\mathcal{E}_0, +)$ est un sous-groupe de $\mathcal{E}$ par les propriétés de la proposition 2.30 (il s'agit du fait que la somme de deux suites convergent vers zéro est une suite convergente vers zéro).

En ce qui concerne la propriété fondamentale des idéaux, si $x \in \mathcal{E}_0$ et $y \in \mathcal{E}$ nous devons prouver que $xy \in \mathcal{E}_0$. Vu que $(\mathcal{E}_0, \cdot\ )$ est commutatif, cela suffira pour être un idéal bilatère. Vu que y est une suite de Cauchy, elle est bornée ; et étant donné que $x \to 0$ nous avons alors $xy \to 0$ (par la proposition 2.30(3)). □

Théorème-définition 2.39 (L'anneau des réels[8])**.**
Sur l'ensemble quotient $\mathcal{E}/\mathcal{E}_0$, les opérations

(1) $\bar{u} + \bar{v} = \overline{u+v}$

(2) $\bar{u} \cdot \bar{v} = \overline{uv}$

sont bien définies et donnent à $\mathcal{E}/\mathcal{E}_0$ une structure de corps commutatif appelé ***corps des réels*** *et noté* $\mathbb{R}$

Démonstration. Nous divisons la preuve en plusieurs parties.

Les opérations sont bien définies Si $u, v \in \mathcal{E}$ et $h, k \in \mathcal{E}_0$ alors $h + k \in \mathcal{E}_0$ et nous avons

$$\overline{u+h} + \overline{v+k} = \overline{u+v+h+k} = \overline{u+v}. \tag{2.47}$$

ainsi que

$$\overline{u+h} \cdot \overline{v+k} = \overline{(u+h)(v+k)} = \overline{uv+uk+hv+hk} = \overline{uv} \tag{2.48}$$

parce que les suites des Cauchy u et v sont bornées, ce qui donne que uk, hv et hk sont des éléments de $\mathcal{E}_0$ par la proposition 2.30(3). Cela prouve que les définitions proposées sont bonnes.

Caractérisation des classes Soit $q \in \mathbb{Q}$ et une suite x convergente vers q. Cette suite est de Cauchy comme toute suite convergente. Montrons que

$$\bar{x} = \{\text{suites qui convergent vers } q\}. \tag{2.49}$$

Si $y \in \bar{x}$ alors $y = x + h$ avec $h \in \mathcal{E}_0$, et comme $h_n \to 0$, on a $y_n \to q$. Réciproquement, si $y_n \to q$ alors pour chaque n nous avons

$$y_n = x_n + (y_n - x_n), \tag{2.50}$$

mais $y_n - x_n \to 0$. Donc la suite $y - x \in \mathcal{E}_0$ ce qui signifie que $y \in \bar{x}$.

Neutre et unité Il est vite vérifié que $\bar{0}$ et $\bar{1}$, les classes des suites constantes sont neutre pour l'addition et unité pour la multiplication.

Corps Nous devons prouver que tout élément non nul est inversible. C'est à dire que si $x \in \mathcal{E}$ ne converge pas vers zéro[14] alors il existe $y \in \mathcal{E}$ tel que $xy \in \bar{1}$.

Nous savons par la proposition 2.30(6) que x étant une suite de Cauchy dans $\mathbb{Q}$, il existe $n_0 \in \mathbb{N}$ tel que $\left(\frac{1}{x_n}\right)_{n \geqslant n_0}$ est une suite de Cauchy. Nous posons alors

$$y_n = \begin{cases} 0 & \text{si } n \leqslant n_0 \\ \frac{1}{x_n} & \text{si } n > n_0. \end{cases} \tag{2.51}$$

13. Définition 2.24.
14. $x \in \mathcal{E}$ peut soit ne pas converger du tout, soit converger vers autre chose que zéro.

Nous avons alors

$$(xy)_n = \begin{cases} 0 & \text{si } n \leqslant n_0 \\ 1 & \text{si } n > n_0 \end{cases} \tag{2.52}$$

et donc $xy \in \bar{1}$.

□

Proposition 2.40.
Soit l'application

$$\begin{aligned} \varphi\colon \mathbb{Q} &\to \mathbb{R} \\ q &\mapsto \bar{q}. \end{aligned} \tag{2.53}$$

où par $\bar{q}$ nous entendons la classe de la suite constante égale à q (qui est de Cauchy).

(1) C'est un homomorphisme injectif.

(2) Image(φ) *est un sous-corps de* $\mathbb{R}$

(3) $\varphi\colon \mathbb{Q} \to$ Image(φ) *est un isomorphisme de corps.*

Démonstration. Le fait que ce soit un homomorphisme est simplement
— $\varphi(q + q') = \overline{q + q'} = \bar{q} + \overline{q'}$
— $\varphi(qq') = \overline{qq'} = \bar{q}\overline{q'}$.
En ce qui concerne l'injectivité, si q est tel que $\varphi(q) = \bar{0} = \mathcal{E}_0$, c'est que

$$\varphi(q) = \{\text{suites de Cauchy qui convergent vers zéro}\} \tag{2.54}$$

Mais nous savons aussi que [15]

$$\varphi(q) = \bar{q} = \{\text{suites de Cauchy qui convergent vers } q\} \tag{2.55}$$

Nous en déduisons que $q = 0$. □

Lorsque dans la suite nous parlerons d'un élément de $\mathbb{Q}$ comme étant un réel, nous aurons en tête l'image de cet élément par φ.

2.6.2 Relation d'ordre

Nous définissons les parties $\mathcal{E}^+$ et $\mathcal{E}^-$ de $\mathcal{E}$ par

(1) $x \in \mathcal{E}^+$ si et seulement si pour tout $\epsilon > 0$, il existe N_ϵ tel que $n > N_\epsilon$ implique $x_n > -\epsilon$.

(2) $x \in \mathcal{E}^-$ si et seulement si pour tout $\epsilon > 0$, il existe N_ϵ tel que $n > N_\epsilon$ implique $x_n < \epsilon$.

Nous notons aussi $\mathcal{E}^{++} = \mathcal{E}^+ \backslash \mathcal{E}_0$.

Lemme 2.41.
Les parties $\mathcal{E}^+$ et $\mathcal{E}^-$ partitionnent $\mathcal{E}$ de la façon suivante :

(1) $\mathcal{E}^+ \cap \mathcal{E}^- = \mathcal{E}_0$

(2) $\mathcal{E}^+ \cup \mathcal{E}^- = \mathcal{E}$

Démonstration. On prouve d'abord que $\mathcal{E}^+ \cap \mathcal{E}^- \subset \mathcal{E}_0$, l'inclusion inverse est évidente. Soit $\epsilon > 0$ et $x \in \mathcal{E}^+ \cap \mathcal{E}^-$. Il existe un $N \in \mathbb{N}$ tel que $x_n > -\epsilon$ et $x_n < \epsilon$ pour tout $n \geqslant N$. Par conséquent, $|x_n| \leqslant \epsilon$ pour tout $n \geqslant N$ et la suite x converge vers zéro, c'est à dire $x \in \mathcal{E}_0$.

Pour prouver le second point, soit $x \in \mathcal{E} \backslash \mathcal{E}^-$, et prouvons que $x \in \mathcal{E}^+$. La condition $x \notin \mathcal{E}^-$ donne qu'il existe un $\alpha > 0$ (dans $\mathbb{Q}$) tel que pour tout n, il existe $p > n$ avec $x_p > \alpha$. Mais x est une suite de Cauchy, donc nous avons un n_0 tel que si $n, p \geqslant n_0$ alors $|x_n - x_p| \leqslant \frac{\alpha}{2}$. En particulier, si $n \geqslant n_0$, et si $p > n$ est tel que $x_p > \alpha$, on obtient

$$x_n > \frac{\alpha}{2} > 0 \tag{2.56}$$

Par conséquent $x \in \mathcal{E}^+$ parce que $x \in \mathcal{E}$ et les x_n sont tous positifs à partir d'un certain rang. □

15. Voir dans la démonstration du théorème 2.39.

Lemme 2.42 ([8])**.**
Quelques propriétés du partitionnement.

(1) $x \in \mathcal{E}^-$ *si et seulement si* $(-x) \in \mathcal{E}^+$

(2) $x \in \mathcal{E}^+$ *et* $y \in \mathcal{E}^+$ *implique* $x + y \in \mathcal{E}^+$

(3) $x \in \mathcal{E}^+$ *et* $y \in \mathcal{E}^+$ *implique* $xy \in \mathcal{E}^+$

(4) Si $x, y \in \mathcal{E}$ *sont tels que* $x - y \in \mathcal{E}_0$ *alors soit* $x, y \in \mathcal{E}^+$ *soit* $x, y \in \mathcal{E}^-$.

Démonstration. Point par point.

(1) Définition de $\mathcal{E}^+$ et $\mathcal{E}^-$.

(2) Pour $n \geqslant N_{\epsilon/2}$ nous avons $x_n > -\epsilon/2$ et $y_n > -\epsilon/2$. Donc $x_n + y_n > -\epsilon$.

(3) Si x ou y est dans $\mathcal{E}_0$ alors $xy \in \mathcal{E}_0$ et c'est bon. Si par contre $x, t \in \mathcal{E}^{++}$ alors pour n suffisamment grand, $x_n > 0$ et $y_n > 0$. Et dans ce cas, $(xy)_n > 0$, c'est à dire $xy \in \mathcal{E}^+$.

(4) Supposons que $x - y \in \mathcal{E}_0$ avec $x \in \mathcal{E}^+$ et prouvons qu'alors $y \in \mathcal{E}^+$. Soit donc $\epsilon > 0$; il existe n_1 tel que $x_n > -\frac{\epsilon}{2}$ dès que $n \geqslant n_1$. Mais $x - y \in \mathcal{E}_0$, donc il existe n_2 tel que $|x_n - y_n| < \frac{\epsilon}{2}$ dès que $n \geqslant n_2$. En prenant n plus grand que n_1 et n_2, nous avons en même temps

$$\begin{cases} x_n > -\dfrac{\epsilon}{2} & \text{(2.57a)} \\ |x_n - y_n| < \dfrac{\epsilon}{2}. & \text{(2.57b)} \end{cases}$$

Cela implique que $y_n > -\epsilon$ et donc que $y \in \mathcal{E}^+$.
Nous pouvons de même prouver que si $x \in \mathcal{E}^-$ alors $y \in \mathcal{E}^-$.

□

Définition 2.43 (Positivité dans $\mathbb{R}$)**.**
Vocabulaire et notations.

(1) Nous notons $\mathbb{R} = \mathcal{E}/\mathcal{E}_0$.

(2) Nous notons $\mathbb{R}^+ = \mathcal{E}^+$.

(3) Nous notons $\mathbb{R}^- = \mathcal{E}^-$.

(4) Un élément de $\mathbb{R}$ *est* ***positif*** *s'il est la classe d'une suite de Cauchy appartenant à* $\mathcal{E}^+$.

(5) Un élément de $\mathbb{R}$ *est* ***négatif*** *s'il est la classe d'une suite de Cauchy appartenant à* $\mathcal{E}^-$.

*(6) Lorsque nous parlons de nombres réels, le symbole «*0*» signifie* $\mathcal{E}_0$ *ou plus précisément la classe d'un élément de* $\mathcal{E}_0$ *modulo* $\mathcal{E}_0$.

Remarque 2.44.
Avec les conventions de la définition 2.43, et en anticipant sur nos connaissances à propos des réels,

(1) zéro est positif et négatif.

(2) L'intersection entre $\mathbb{R}^+$ et $\mathbb{R}^-$ est le singleton $\{0\}$.

(3) L'ensemble des nombres *strictement* positifs est noté $(\mathbb{R}^+)^*$ ou $\mathbb{R}^+\backslash\{0\}$.

(4) Le mot «positif» signifie «positif ou nul» ; le mot «négatif» signifie «négatif ou nul».

Problèmes et choses à faire

Une inconsistance dans les notations et le vocabulaire en ce qui concerne $\mathbb{R}^+$, $\mathbb{R}^-$ et le fait que zéro soit dedans ou non m'a été signalée. Nous avons tranché pour les conventions données dans la remarque 2.44 qui sont aussi celles de Wikipédia, mais il n'est pas exclu qu'il y ait des incohérences par-ci par-là dans le texte. Merci de me les signaler.

Définition 2.45 (Ordre dans $\mathbb{R}$)**.**
Si $a, b \in \mathbb{R}$ *nous notons* $a \leqslant b$ *si et seulement si* $b - a$ *est positif. Nous notons aussi* $a > b$ *si et seulement si* $b - a \in \mathbb{R}^+\backslash\{0\}$, *etc.*

Lemme 2.46.
Les premières propriétés de l'ordre.

(1) L'ensemble $(\mathbb{R}, \leqslant)$ *est un corps totalement ordonné (définitions 2.4 et 2.25).*

(2) L'application

$$\begin{aligned} \varphi\colon \mathbb{Q} &\to \mathbb{R} \\ q &\mapsto \bar{q} \end{aligned} \tag{2.58}$$

dont nous avons déjà parlé dans la proposition 2.40 est strictement croissante.

Démonstration. Nous prouvons la stricte croissance de φ. Si $q < l$ alors $\varphi(q) - \varphi(l) = \overline{q-l}$ est la classe de la suite constante $q - l$ qui est un élément strictement positif de $\mathbb{Q}$. Nous avons donc $\overline{q-l} \in \mathbb{R}^+$, et donc $\varphi(q) < \varphi(l)$. □

Remarque 2.47.
Comme déjà mentionné plus haut, à chaque fois que nous parlerons d'un élément de $\mathbb{Q}$ comme étant un élément de $\mathbb{R}$, nous considérons la classe de la suite constante.

Lemme 2.48.
Si $x, y, z \in \mathbb{R}$ *avec* $x > 0$ *sont tels que* $z > y/x$ *alors* $zx > y$.

Démonstration. Nous savons que

$$z - \frac{y}{x} \in \mathcal{E}^+ \backslash \{0\} = \mathcal{E}^{++}. \tag{2.59}$$

Vu que $x \in \mathcal{E}^{++}$, multiplier par x fait rester dans $\mathcal{E}^{++}$:

$$zx - x\frac{y}{x} \in \mathcal{E}^{++}. \tag{2.60}$$

Un représentant de $x\frac{y}{x}$ est la suite $n \mapsto x_n \frac{y_n}{x_n} = y_n$. Donc $x\frac{y}{x} = y$. Cela signifie que $zx - y \in \mathcal{E}^{++}$ et donc que $zx > y$. □

Lemme 2.49.
Pour tout $a \in \mathbb{R}$*, il existe* $p \in \mathbb{N}$ *tel que* $p > a$.

Démonstration. Nous allons donner deux preuves différentes de ce lemme.

Première façon L'élément a de $\mathbb{R}$ admet un représentant (a_n) qui est une suite de Cauchy dans $\mathbb{Q}$. C'est donc une suite bornée, c'est-à-dire qu'il existe $m, q \in \mathbb{N}$ tels que $|a_n| \leqslant m/q$ pour tout n (proposition 2.30(2)). Soit M un naturel strictement plus grand que m/q [16].

La suite de Cauchy $(M - a_n)_{n\in\mathbb{N}}$ est constituée de rationnels positifs et est donc dans $\mathcal{E}^+$. La classe de $M - a$ est donc un réel positif [17]. Par définition de la relation d'ordre, $M \geqslant a$.

Seconde façon La suite (a_n) est majorée par $\frac{m}{q}$, donc on a dans $\mathbb{Q}$ et pour tout n :

$$a_n \leqslant \frac{m}{q} = M \leqslant qM. \tag{2.61}$$

L'application $\varphi\colon \mathbb{Q} \to \mathbb{R}$ est croissante, donc

$$\varphi\big((a_n)\big) \leqslant \varphi(qM). \tag{2.62}$$

□

En corollaire, nous avons

Lemme 2.50.
Pour tout $x \in \mathbb{R}$*, il existe* $q \in \mathbb{Z}$ *tel que* $q < x$.

Démonstration. Utilisation du lemme précédent avec $a = -x$: on prend $q = -p$. □

16. Lemme 2.29.
17. Et nous allons d'ailleurs arrêter de toujours préciser «la classe de» lorsque ce n'est pas nécessaire.

Théorème 2.51 ([8]).
Le corps $\mathbb{R}$ est archimédien[18].

Démonstration. Le lemme 2.46 dit que $\mathbb{R}$ est totalement ordonné. Soient $x, y \in \mathbb{R}$ avec $x > 0$; posons $a = \frac{y}{x}$. Le lemme 2.49 nous donne un $p \in \mathbb{N}$ tel que $p > a$.Nous concluons alors avec le lemme 2.48 :

$$px > ax = \frac{y}{x}x = y. \tag{2.63}$$

□

Le lemme suivant n'est pas loin de dire que $\mathbb{Q}$ est dense dans $\mathbb{R}$, à part que nous n'avons pas encore donné de topologie sur $\mathbb{R}$.

Lemme 2.52.
Si $x, y \in \mathbb{R}$ sont tels que $x < y$, alors il existe $s \in \mathbb{Q}$ tel que $x < s < y$.

Démonstration. Nous avons par hypothèse que $y - x > 0$ et donc le fait que $\mathbb{R}$ soit archimédien (théorème 2.51) nous donne $q \in \mathbb{N}$ tel que $q(y - x) > 1$. Soit

$$E = \{n \in \mathbb{Z} \text{ tel que } \frac{n}{q} \leqslant x\}. \tag{2.64}$$

Cet ensemble n'est pas vide à cause du lemme 2.50 ; de plus, comme $|x|q \leqslant n_0$ pour un certain n_0 (à cause du lemme 2.49), l'ensemble E est majoré par n_0. Donc E possède un plus grand élément [19] p qui vérifie

$$\frac{p}{q} \leqslant x < \frac{p+1}{q}. \tag{2.65}$$

De plus $(p + 1)/q < y$. En effet nous avons

$$\frac{p+1}{q} = \frac{p}{q} + \frac{1}{q} \leqslant x + \frac{1}{q} < x + y - x = y \tag{2.66}$$

où nous avons utilisé l'inégalité stricte $y - x > \frac{1}{q}$.

Nous avons donc

$$x < \frac{p+1}{q} < y, \tag{2.67}$$

et le nombre $(p + 1)/q$ convient comme s. □

Remarque 2.53.
Le lemme 2.52 a également pour conséquence que des ensembles comme $[-1, 1]$ ne sont pas bien ordonnés (définition 2.4). En effet la partie $]0, 1[$ ne possède pas de minimum parce que si $x \in]0, 1[$ alors $0 < x$ et il existe $s \in \mathbb{Q}$ (a fortiori $s \in \mathbb{R}$) tel que $0 < s < x$, c'est à dire que x n'est pas un minimum de $]0, 1[$.

Nous avons terminé avec la construction des réels. Les propriétés topologiques arrivent en 6.4. En particulier, le fait que $\mathbb{R}$ soit complet est le théorème 6.57.

2.7 Les complexes

Problèmes et choses à faire

Encore une fois, cette section n'est pas du tout faite. Le but de cette section serait de

- Construire $\mathbb{C}$ en tant qu'ensemble
- Construire les opérations courantes.
- Démontrer les propriétés de base.

Attention : il est sans espoir de parler de forme trigonométrique ici parce que les exponentielles et fonctions trigonométriques ne sont définies qu'avec les séries.

Nous donnons ici en vrac quelque propriétés et définitions que se doivent d'être présentes dans la version finale de cette section, si elle existe un jour.

Comme partout dans ce chapitre sur la construction des ensembles de nombres, certains propriétés qui ont l'air toute simples peuvent, en fonction des définitions prises, s'avérer pas du tout évidentes.

18. Définition 2.25(5).
19. Lemme 2.19.

2.7.1 Définitions

Un nombre complexe s'écrit sous la forme $z = a + bi$, où a et b sont des nombres réels appelés (et notés) respectivement partie réelle ($a = \Re(z)$) et partie imaginaire ($b = \Im(z)$) de z. L'ensemble des nombres de cette forme s'appelle l'ensemble des nombres complexes ; cet ensemble porte une structure de corps et est noté $\mathbb{C}$. Le nombre complexe $i = 0 + 1i$ est un nombre imaginaire qui a la particularité que $i^2 = -1$.

Deux nombres complexes $a+bi$ et $c+di$ sont égaux si et seulement si $a = c$ et $b = d$, c'est-à-dire leurs parties réelles sont égales, et leurs parties imaginaires sont égales.

Pour $z = a + bi$ un nombre complexe, on note $\bar{z} = a - bi$ le *complexe conjugué* de z. Dans le plan de Gauss, il s'agit du symétrique de z par rapport à la droite réelle (généralement dessinée horizontalement).

On définit le module du complexe z par $|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$. Dans le plan de Gauss, il s'agit de la distance entre 0 et z.

Proposition 2.54.
Pour tous $z = a + bi$ et z' nombres complexes, on a

(1) $z\bar{z} = a^2 + b^2$;

(2) $\bar{\bar{z}} = z$;

(3) $|z| = |\bar{z}|$;

(4) $|zz'| = |z|\,|z'|$;

(5) $|z + z'| \leqslant |z| + |z'|$.

Lemme 2.55.
Pour tout $z \in \mathbb{C}$ nous avons $z\bar{z} = \bar{z}z = |z|^2$.

Chapitre 3

Théorie des groupes

3.1 Groupes

Définition 3.1 ([11])**.**
Soit G un groupe. Le ***centralisateur*** *de $H \subset G$ est*

$$\mathcal{Z}_G(H) = \{g \in G \text{ tel que } hg = gh \, \forall h \in H\}; \tag{3.1}$$

il contient donc tous les éléments de G qui commutent avec ceux de H.

Si H est un sous-groupe, son ***normalisateur*** *est*

$$N_G(H) = \{g \in G \text{ tel que } gH = Hg\}. \tag{3.2}$$

Définition 3.2.
Le ***centre*** *d'un groupe G est l'ensemble des éléments de G qui commutent avec tous les autres :*

$$Z_G = \{z \in G \text{ tel que } gz = zg \forall g \in G\}. \tag{3.3}$$

Si $g \in G$ nous notons $Z_G(g)$ le ***centralisateur*** *de g dans G :*

$$Z_G(g) = \{h \in G \text{ tel que } hg = hg\}. \tag{3.4}$$

C'est l'ensemble des éléments de G qui commutent avec g.

Définition 3.3.
Un sous-groupe N de G est ***normal*** *ou* ***distingué*** *si pour tout $g \in G$ et pour tout $n \in N$, $gng^{-1} \in N$. Autrement dit lorsque $gNg^{-1} \subset N$.*

Lorsque N est normal dans G il est parfois noté $N \lhd G$.

Définition 3.4.
Un sous-groupe H de G est un sous-groupe ***caractéristique*** *si $\alpha(H) = H$ pour tout $\alpha \in \mathrm{Aut}(G)$.*

Définition 3.5 (Groupe simple)**.**
Un groupe est dit ***simple*** *si il est non trivial et si les seuls sous-groupes normaux qu'il admet sont lui-même et le sous-groupe réduit à l'identité.*

Définition 3.6.
Soit G, un groupe et A une partie de G. Nous notons $\mathrm{gr}(A)$ l'intersection de tous les sous-groupes de G contenant A. C'est le plus petit (pour l'inclusion) groupe de G contenant A. Si $\mathrm{gr}(A) = G$, alors nous disons que A est une ***partie génératrice*** *le groupe G.*

Un groupe est ***monogène*** *s'il a une partie génératrice réduite à un seul élément.*

Définition 3.7 (Groupe cyclique)**.**
Un élément $a \in G$ est un ***générateur*** *de G si tous les éléments de G s'écrivent sous la forme a^n pour un certain $n \in \mathbb{Z}$. Un groupe fini et monogène est dit* ***cyclique****.*

Définition 3.8 (Sous-groupe engendré)**.**
*Soit A une partie du groupe G. Le sous-groupe **engendré** est l'intersection de tous les sous-groupes de G contenant A.*

Lemme 3.9.
Si A est une partie du groupe G, alors le sous-groupe engendré[1] *par A est l'ensemble de tous les produits finis d'éléments de A et de A^{-1} (l'identité est le produit à zéro éléments).*

Démonstration. Nous nommons K le groupe engendré par A et par H l'ensemble de tous les produits finis d'éléments de A et de A^{-1}. Cela est un sous-groupe de G contenant A. Donc $K \subset H$ parce que K est une intersection dont un des éléments est H.

Par ailleurs tout groupe contenant A doit contenir les inverses et les produits finis, donc $H \subset K$.
Au final, $H = K$, ce qu'il fallait. □

Exemple 3.10
Soit G le groupe des rotations d'angle $2k\pi/10$. La rotation d'angle $\pi/2$ n'est pas génératrice parce qu'elle n'engendre que $\pi/2$, π,$3\pi/2$ et l'identité. La rotation d'angle $2\pi/10$ par contre est génératrice. △

3.2 Sous groupe normal

Proposition 3.11.
Soit N un sous-groupe de G. Les propriétés suivantes sont équivalentes :

(1) $gNg^{-1} \subseteq N$ pour tout $g \in G$,
(2) $gNg^{-1} = N$ pour tout $g \in G$,
(3) $gN = Ng$ pour tout $g \in G$,
(4) N est une union de classes de conjugaison de G,
(5) N est normal.

Définition 3.12.
Soit $g \in G$ et $n \in \mathbb{Z}$. Nous définissons g^n par

(1) $g^0 = e$ et $g^n = gg^{n-1}$ si n est positif.
(2) si $n < 0$, nous posons $g^n = (g^{-1})^{-n}$.

Définition 3.13 (Ordre d'un groupe et d'un élément)**.**
Ce sont deux choses différentes.

*(1) Si G est un groupe, l'**ordre** est la cardinalité de G et est noté $|G|$.*
*(2) L'**ordre** d'un élément g de G est le naturel*

$$\min\{n \in \mathbb{N} \text{ tel que } g^n = e\}, \tag{3.5}$$

s'il existe ; dans le cas contraire, nous disons que l'ordre de g est infini.

Nous verrons que le corollaire 3.55 au théorème de Lagrange dira que l'ordre d'un élément divise l'ordre du groupe.

Lemme 3.14 ([12])**.**
Si H et K sont normaux dans le groupe G et si $H \cap K = \{e\}$ alors $HK \simeq H \times K$.

Définition 3.15.
*L'**exposant** du groupe G est le plus petit entier non nul n tel que $g^n = e$ pour tout $g \in G$. S'il n'existe pas, nous disons que l'exposant du groupe est infini.*

1. Définition 3.8.

Si l'ordre de tous les éléments acceptent un majorant commun, alors l'exposant du groupe est le plus petit commun multiple des ordres des éléments. En particulier pour un groupe fini, l'exposant est le ppcm des ordres des éléments du groupe.

Le théorème de Burnside 9.238 nous donnera un bon paquet d'exemples de groupes d'exposant fini dans $\mathrm{GL}(n, \mathbb{C})$.

Proposition 3.16.
Soit H un sous-groupe normal de G et $\psi: G \to K$ un homomorphisme.

(1) $\psi(H)$ est normal dans $\psi(G)$

(2) Si G/H est abélien alors $\psi(G)/\psi(H)$ est abélien.

Démonstration. Soient $h \in H$ et $g \in G$. Alors $\psi(g)\psi(h)\psi(g)^{-1} = \psi(ghg^{-1}) \in \psi(H)$. Donc $\psi(H)$ est normal dans $\psi(G)$.

Pour la seconde partie nous notons $[\dots]$ les classes par rapport à $\psi(H)$ et $\overline{\dots}$ celles par rapport à H. Nous avons

$$\begin{aligned}
[\psi(g_1)][\psi(g_2)] &= \big[\psi(g_1)\psi(g_2)\big] && (3.6a)\\
&= \big[\psi(g_1g_2)\big] && (3.6b)\\
&= \{\psi(g_1g_2)\psi(h) \text{ tel que } h \in H\} && (3.6c)\\
&= \{\psi(g_1g_2h) \text{ tel que } h \in H\} && (3.6d)\\
&= \psi\Big(\{g_1g_2h \text{ tel que } h \in H\}\Big) && (3.6e)\\
&= \psi\big(\overline{g_1g_2}\big) && (3.6f)\\
&= \psi(\overline{g_2g_1}) && (3.6g)\\
&= \text{refaire à l'envers} && (3.6h)\\
&= [\psi(g_2)][\psi(g_1)]. && (3.6i)
\end{aligned}$$

Par conséquent $\psi(G)/\psi(H)$ est abélien. □

3.2.1 Classes de conjugaison

Définition 3.17.
*Soit un groupe G et un élément $g \in G$. La **classe de conjugaison** de g est la partie*

$$C_g = \{kgk^{-1} \text{ tel que } k \in G\}. \tag{3.7}$$

Lemme 3.18.
Un groupe est abélien si et seulement si ses classes de conjugaisons sont des singletons.

Démonstration. Supposons que G soit abélien. Alors

$$C_g = \{kgk^{-1} \text{ tel que } k \in G\} = \{g\}. \tag{3.8}$$

Donc les classes de conjugaison sont des singletons.

Dans l'autre sens, si les classes sont des singletons, on a $kgk^{-1} = g$ pour tout $k, g \in G$. Cela signifie immédiatement que G est abélien. □

3.3 Groupe dérivé

Définition 3.19.
*Si G est un groupe et si $g, h \in G$, nous notons $[g, h] = ghg^{-1}h^{-1}$ le **commutateur** de g et h. Le **groupe dérivé** de G est le sous-groupe noté $D(G)$ ou $[G, G]$ engendré par les commutateurs.*

Autrement dit, $D(G)$ est l'intersection de tous les sous-groupes de G contenant tous les commutateurs. Intersection non vide parce que G lui-même en fait partie.

En vertu du lemme 3.9, le groupe dérivé de G est l'ensemble des produits finis de commutateurs. C'est à dire que si S_m est l'ensemble des produits de m commutateurs, alors

$$D(G) = \bigcup_{m=1}^{\infty} S_m. \tag{3.9}$$

Lemme 3.20.
Le groupe dérivé est un sous-groupe caractéristique, et un sous-groupe normal.

Démonstration. Il est évident que si $\alpha \in \mathrm{Aut}(G)$ alors

$$\alpha\big([g,h]\big) = \big[\alpha(g), \alpha(h)\big], \tag{3.10}$$

c'est à dire que $D(G)$ est un sous-groupe caractéristique. En particulier si c est un commutateur, alors xcx^{-1} en est encore un, ce qui montre que $D(G)$ est normal dans G. Plus spécifiquement,

$$x(ghg^{-1}h^{-1})x^{-1} = (xgx^{-1})(xhx^{-1})(xg^{-1}x^{-1})(xh^{-1}x^{-1}) \tag{3.11a}$$
$$= (xgx^{-1})(xhx^{-1})(xgx^{-1})^{-1}(xhx^{-1})^{-1}. \tag{3.11b}$$

□

Proposition 3.21.
Le groupe quotient $G/D(G)$ est abélien.

Démonstration. En ce qui concerne le fait que $G/D(G)$ soit abélien, nous savons que pour tout $g, h \in G$ nous avons $h^{-1}g^{-1}hg \in D(G)$ et donc

$$[g][h] = [gh] = [ghh^{-1}g^{-1}hg] = [hg] = [h][g]. \tag{3.12}$$

□

Le groupe quotient $G/D(G)$ est appelé l'**abélianisé** de G et est parfois noté G^{ab}.

Si $f\colon G \to A$ est un homomorphisme entre le groupe G et un groupe abélien A, alors $f\big(D(G)\big) = \{0\}$. Du coup f passe au quotient de G par $D(G)$, et il existe une unique application $\bar{f}\colon G/D(G) \to A$ telle que $f = \bar{f} \circ \pi$ où $\pi\colon G \to G/D(G)$ est la projection canonique.

3.4 Théorèmes d'isomorphismes

Si G est un groupe et si N est un sous-groupe normal, alors l'ensemble G/N a une structure de groupe et la projection canonique $\pi\colon G \to G/N$ est un homomorphisme surjectif de noyau N.

Théorème 3.22 (Premier théorème d'isomorphisme)**.**
Soit $\theta\colon G \to H$ un homomorphisme de groupe. Alors

(1) $\mathrm{Ker}\,\theta$ *est normal dans G,*

(2) $\mathrm{Image}\,\theta$ *est un sous-groupe de H*

(3) nous avons un isomorphisme naturel

$$G/\mathrm{Ker}\,\theta \simeq \mathrm{Image}\,\theta \tag{3.13}$$

Démonstration. Point par point.

(1) Le fait que $\mathrm{Ker}\,\theta$ soit un sous-groupe de G est clair ; montrons qu'il est normal. Si $g \in G$ et $u \in \mathrm{Ker}\,\theta$, alors $\theta(g^{-1}ug) = \theta(g^{-1})\theta(u)\theta(g) = \big(\theta(g)\big)^{-1}\theta(g) = 1_H$, et donc $g^{-1}ug \in \mathrm{Ker}\,\theta$.

(2) Il suffit de remarquer que si $h = \theta(g)$ et $h' = \theta(g')$, alors $h^{-1}h' = \theta(g^{-1}g')$.

(3) Si $[g]$ représente la classe de g dans $G/\operatorname{Ker}\theta$, l'isomorphisme est donné par $\varphi[g] = \theta(g)$.

□

Théorème 3.23 (Deuxième théorème d'isomorphisme)**.**
Soient H et N deux sous-groupes de G et supposons que N soit normal. Alors

(1) $NH = HN$ est un sous-groupe.

(2) Le groupe N est normal dans NH.

(3) Le groupe $N \cap H$ est normal dans H.

(4) Nous avons l'isomorphisme

$$\frac{HN}{N} \simeq \frac{H}{H \cap N}. \tag{3.14}$$

(5) L'isomorphisme du point (4) est encore valable si N n'est pas normal mais si seulement H normalise N, c'est à dire si $hNh^{-1} \in N$ pour tout $h \in H$.

Démonstration. Point par point.

(1) Il est clair que $1_G \in NH$. Soient nh et $n'h'$ deux éléments de NH ; alors en tenant compte du fait que N est normal,

$$nhn'h' = n\underbrace{hn'h^{-1}}_{\in N}hh' \in NH. \tag{3.15}$$

Cela prouve que NH est un groupe.

De la même façon, nous prouvons que HN est un groupe par

$$hnh'n' = hh'\underbrace{h'^{-1}nh'}_{\in N}n' \in HN \tag{3.16}$$

Nous devons encore prouver que $HN = NH$. Pour cela, $nh \in HN$, car $nh = hh^{-1}nh$, les trois derniers facteurs formant un élément de N par normalité ; de même $hn \in NH$, montrant que $NH = HN$. Enfin, comme $(nh)^{-1} = h^{-1}n^{-1}$), les inverses de NH sont dans $HN = NH$.

(2) N est normal dans G, a fortiori dans l'un de ses sous-groupes.

(3) Il suffit de voir que, si $h \in H$ et $n \in N \cap H$, alors $hnh^{-1} \in N \cap H$. Or, $hnh^{-1} \in H$ puisque H est un sous-groupe ; et $hnh^{-1} \in N$ car N est un sous-groupe normal de G.

(4) Il faut d'abord remarquer que H et N étant des groupes et le produit NH étant un groupe, nous avons $NH = HN$. Soit le morphisme injectif

$$\begin{aligned} j\colon H &\to HN \\ h &\mapsto h \end{aligned} \tag{3.17}$$

et la surjection canonique

$$\sigma\colon HN \to HN/N \tag{3.18}$$

Nous considérons ensuite l'application composée

$$\begin{aligned} f\colon H &\to HN/N \\ h &\mapsto hN. \end{aligned} \tag{3.19}$$

$\underline{f \text{ **est surjective**}}$ L'application f est surjective parce que l'élément $hnN \in HN/N$ est l'image de h, étant donné que $hnN = hN$.

$\underline{\operatorname{Ker}(f) = H \cap N}$ Si $a \in H \cap N$, nous avons $f(a) = aN = N$, et donc $H \cap N \subset \operatorname{Ker} f$. D'autre part, si $h \in H$ vérifie $h \in \operatorname{Ker} f$, alors $f(h) = hN = N$, ce qui est uniquement possible si $h \in N$.

Le premier théorème d'isomorphisme implique alors que $H/\operatorname{Ker} f \simeq \operatorname{Image} f$, c'est à dire

$$H/N \cap H \simeq HN/N. \tag{3.20}$$

□

Théorème 3.24 (Troisième théorème d'isomorphisme)**.**
Soient N et M deux sous-groupes normaux de G avec $M \subset N$. Alors N/M est normal dans G/M et

$$(G/M)/(N/M) \simeq G/N. \tag{3.21}$$

Démonstration. Afin de montrer que N/M est normal dans G/M, nous considérons $g \in G$, $nM \in N/M$ et nous calculons

$$gnMg^{-1} = gng^{-1}\underbrace{gMg^{-1}}_{=M} = \underbrace{gng^{-1}}_{\in N}M \in N/M. \tag{3.22}$$

Pour prouver l'isomorphisme nous considérons le morphisme

$$\begin{aligned} \varphi\colon G/M &\to G/N \\ gM &\mapsto gN. \end{aligned} \tag{3.23}$$

C'est surjectif et le noyau est N/M parce que $\varphi(gM) = N$ uniquement si $g \in N$. Nous pouvons appliquer le premier théorème d'isomorphisme à φ en écrivant

$$(G/M)/\operatorname{Ker}\varphi \simeq \operatorname{Image}\varphi, \tag{3.24}$$

c'est à dire

$$(G/M)/(N/M) \simeq G/N. \tag{3.25}$$

□

3.5 Le groupe et anneau des entiers

Certes $\mathbb{Z}$ est un groupe pour l'addition, mais c'est également un anneau [2] parce que nous avons les deux opérations d'addition et de multiplication. Nous n'allons pas nous priver de cette belle structure juste parce que le titre du chapitre est «groupes».

3.5.1 Division euclidienne

Théorème 3.25 (Division euclidienne[13])**.**
Soient $a \in \mathbb{Z}$ et $b \in \mathbb{N}^*$. Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$, avec $0 \leqslant r < b$, tel que

$$a = bq + r. \tag{3.26}$$

Démonstration. Remarquons que $r = a - bq$, et donc, une fois l'existence et l'unicité de q établie, celle de r suivra.

Unicité Nous supposons avoir $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que $0 \leqslant r < b$ et $a = bq + r$. Alors forcément $r = a - qb$ et $0 \leqslant a - qb < b$, ou encore

$$qb \leqslant a < (1 + q)b. \tag{3.27}$$

Ces deux inéquations fixent $q \in \mathbb{Z}$. En effet nous démontrons maintenant que seul $k = 0$ permet à $q + k$ de vérifier ces deux inéquations (parmi les $k \in \mathbb{Z}$). Si $g \geqslant 1$ alors

$$(q + k)b = (q + 1)b + (k - 1)b > a. \tag{3.28}$$

Et si $k \leqslant -1$ alors

$$(q + k + 1)b \leqslant qb \leqslant a. \tag{3.29}$$

D'où l'unicité de q et par conséquent celle de r.

2. Définition 2.20.

Existence Nous considérons l'ensemble

$$E = \{q \in \mathbb{Z} | bq \leqslant a\}.$$

C'est un sous-ensemble d'entiers non-vide (il contient $-|a|$) et admet $|a|$ comme majorant ; il admet donc un maximum q. Ce maximum vérifie

$$bq \leqslant a < b(q+1). \tag{3.30}$$

Cela donne $0 \leqslant a - bq < b$ et le résultat en posant $r = a - qb$.

□

Définition 3.26.
*L'opération $(a, b) \mapsto (q, r)$ donnée par le théorème 3.25 est la **division euclidienne**. Le nombre q est le **quotient** et r est le **reste** de la division de a par b.*

3.5.2 Sous-groupes de $(\mathbb{Z}, +)$

Proposition 3.27.
Une partie H du groupe $(\mathbb{Z}, +)$ est un sous-groupe si et seulement s'il existe $n \in \mathbb{N}$ tel que $H = n\mathbb{Z}$.

Démonstration. Soit $H \neq \{0\}$ un sous-groupe de $\mathbb{Z}$. L'ensemble $H \cap \mathbb{N}^*$ contient un élément minimum que nous notons n. Nous avons certainement $n\mathbb{Z} \subset H$ parce que H est un groupe (donc $n + n$ et $-n$ sont dans H dès que n est dans H). Nous devons prouver que $H \subset n\mathbb{Z}$.

Si $x \in H$, par le théorème de division euclidienne 3.25, il existe $q \in \mathbb{Z}$ et $r \in \mathbb{N}$, uniques, tels que $x = nq + r$ et $0 \leqslant r < n$. Nous savons déjà que $nq \in H$, donc $r = x - nq \in H$. Le nombre r est donc un élément de H strictement plus petit que n. Mais nous avions décidé que n serait le plus petit élément de $H \cap \mathbb{N}^*$. Par conséquent $r = 0$ et $x = nq \in n\mathbb{Z}$. □

Notons que si un sous-groupe H de $\mathbb{Z}$ est donné, alors le nombre n tel que $H = n\mathbb{Z}$ est unique. En effet si $n\mathbb{Z} = m\mathbb{Z}$ nous avons que n divise m (parce que $m \in m\mathbb{Z} \subset n\mathbb{Z}$) et que m divise n parce que $n \in m\mathbb{Z}$. Par conséquent $n = m$.

3.5.3 PGCD, PPCM et Bézout

Voir aussi les définitions de PGCD et PPCM dans un anneau intègre à la définition 4.50.

Définition 3.28 (PPCM et PGCD)**.**
*Soient $p, q \in \mathbb{Z}^*$. Le **pgcd** de p et q est le plus grand diviseur commun de p et q. Le **ppcm** de p et q est leur plus petit multiple commun.*

Lemme 3.29.
Soient $p, q \in \mathbb{Z}^*$. Les entiers $\operatorname{ppcm}(p, q)$ et $\operatorname{pgcd}(p, q)$ fournissent les isomorphismes de groupes suivants :

$$p\mathbb{Z} \cap q\mathbb{Z} = \operatorname{ppcm}(p, q)\mathbb{Z} \tag{3.31a}$$
$$p\mathbb{Z} + q\mathbb{Z} = \operatorname{pgcd}(p, q)\mathbb{Z}. \tag{3.31b}$$

Définition 3.30.
*Si $\operatorname{pgcd}(p, q) = 1$, nous disons que p et q sont **premiers entre eux**. Si nous avons un ensemble d'entiers a_i, nous disons qu'ils sont premiers **dans leur ensemble** si 1 est le PGCD de tous les a_i ensemble.*

Les nombres 2, 4 et 7 ne sont pas premiers deux à deux (à cause de 2 et 4), mais ils sont premiers dans leur ensemble parce qu'il n'y a pas de diviseurs communs à tout le monde.

Définition 3.31.
*Un **nombre premier** est un naturel acceptant exactement deux diviseurs distincts.*

Avec cette définition, 0 n'est pas premier, 1 n'est pas premier et 2 est premier.

Théorème 3.32 (Théorème de Bézout [3][14], thème 47)**.**
Deux entiers non nuls $a, b \in \mathbb{Z}^*$ sont premiers entre eux si et seulement s'il existe $u, v \in \mathbb{Z}$ tels que

$$au + bv = 1 \tag{3.32}$$

Démonstration. Soit $d = \text{pgcd}(a, b)$ et des nombres u, v tels que $au + bv = 1$. Le PGCD d divise à la fois a et b, et donc divise $au + bv$. Nous en déduisons que d divise 1 et est par conséquent égal à 1.

Nous supposons maintenant que $\text{pgcd}(a, b) = 1$ et nous considérons l'ensemble

$$E = \{au + bv \text{ tel que } u, v \in \mathbb{Z}\} \cap \mathbb{N}^*. \tag{3.33}$$

C'est à dire l'ensemble des nombres strictement positifs pouvant s'écrire sous la forme $au + bv$. Cet ensemble est non vide parce qu'il contient par exemple soit a soit $-a$. Soit m le plus petit élément de E et écrivons

$$m = au_1 + bv_1. \tag{3.34}$$

Par le théorème de division euclidienne [4] (avec a et m), il existe des entiers uniques q et r tels que

$$a = mq + r \tag{3.35}$$

avec $0 \leqslant r < m$. En remplaçant m par sa valeur (3.34), $a = (au_1 + bv_1)q + r$ et

$$r = a(1 - u_1 q) - bv_1 q, \tag{3.36}$$

c'est à dire que $r \in \mathbb{Z}a + \mathbb{Z}b$ en même temps que $0 \leqslant r < m$. Si r était strictement positif, il serait dans E. Mais cela est impossible par minimalité de m. Donc $r = 0$ et a est divisible par m.

De la même façon nous prouvons que b est divisible par m. Vu que m divise à la fois a et b nous avons $m = 1$. □

Corollaire 3.33.
Soient p et q deux entiers premiers entre eux. Alors

$$p\mathbb{Z} + q\mathbb{Z} = \mathbb{Z}; \tag{3.37}$$

en particulier, pour tout $x \in \mathbb{Z}$, il existe u_x, v_x entiers tels que $u_x p + v_x q = x$.

Notons que l'application $p\mathbb{Z} + q\mathbb{Z}$ vers $\mathbb{Z}$ n'est évidemment pas injective : les u_x et v_x ne sont pas uniques à x fixé.

Démonstration. Soit $x \in \mathbb{Z}$. Le théorème de Bézout nous donne k et l tels que $kp + lq = 1$. Du coup, $(xk)p + (xl)q = x$. □

La proposition suivante établit que si x est assez grand, alors il peut même être écrit comme une combinaison de p et q à coefficients positifs. Elle sera utilisée pour démontrer que les états apériodiques d'une chaîne de Markov peuvent être atteins à tout moment (assez grand), voir la définition 36.41 et ce qui suit.

Proposition 3.34.
Soient a et b deux éléments de $\mathbb{N}$ premiers entre eux. Il existe $N > 0$ tel que tout $x > N$ appartient à $a\mathbb{N} + b\mathbb{N}$.

Démonstration. Soient a et b, premiers entre eux, et $x \in \mathbb{N}$. Disons tout de suite, pour éviter les cas triviaux et pénibles, que x, a et b sont strictement posifits.

3. Il y a une super application ici : `https://perso.univ-rennes1.fr/matthieu.romagny/agreg/dvt/mauvais_prix.pdf`.

4. Théorème 3.25.

Une décomposition pour x On applique le théorème 3.25 de division euclidienne à x et $a+b$: il existe des entiers p_x, r_x, uniques, tels que

$$\begin{cases} x = (p_x - 1)(a + b) + r_x & (3.38a) \\ 0 \leqslant r_x < a + b. & (3.38b) \end{cases}$$

En d'autres termes, $p_x(a+b)$ est le premier multiple de $a+b$ supérieur ou égal à x. De plus, p_x est strictement positif car x l'est. Il existe alors des entiers u et v tels que

$$ua + vb = p_x(a + b) - x \tag{3.39}$$

par le corollaire 3.33. Ainsi, x peut s'écrire

$$x = (p_x - u)a + (p_x - v)b. \tag{3.40}$$

Des maximums Il s'agit maintenant de savoir si nous pouvons être assuré d'avoir $p_x > u$ et $p_x > v$ dès que x est assez grand. Pour cela, grâce au corollaire 3.33, nous considérons les nombres u_i et v_i définis par

$$u_i a + v_i b = i \tag{3.41}$$

pour $i = 1, \ldots, a+b$. Nous posons $u^* = \max\{u_i\}$, $v^* = \max\{v_i\}$, et $p^* = \max\{u^*, v^*\}$. Nous posons alors $N = p^*(a + b)$, et considérons $x > N$.

Nouvelle décomposition pour x Nous voulons écrire

$$x = (p_x - u_k)a + (p_x - v_k)b \tag{3.42}$$

pour un certain k. Cela demande $u_k a + v_k b = ua + vb = p_x(a+b) - x$ par l'équation (3.39). Vu que $p_x(a+b)-x > 0$, les nombres u_k et v_k existent : il suffit de prendre $k = p_x(a+b)-x$.

Conclusion Avec tous ces choix, nous avons d'abord $x > p^*(a + b)$ et donc

$$x = (p_x - 1)(a + b) + r_x > p^*(a + b), \tag{3.43}$$

ce qui donne

$$(p_x - 1)(a + b) > p^*(a + b) - r_x > (p - 1)(a + b). \tag{3.44}$$

ou encore $p_x > p^*$. Nous avons finalement

$$p_x \geqslant p^* \geqslant u^* \geqslant u_k \tag{3.45}$$

et

$$p_x \geqslant p^* \geqslant v^* \geqslant v_k. \tag{3.46}$$

De ce fait, la décomposition (3.42) est celle que nous voulions.

□

Remarque 3.35.
Une méthode pour obtenir les entiers naturels u et v qui permettent la décomposition $x = au + bv$ est d'abord de choisir u_0 et v_0 tels que au_0 et bv_0 soient les plus proches possibles de $x/2$, puis de décomposer le nombre (relativement petit) $x - au_0 - bv_0$ en $au_1 + bv_1$. Deux nombres u et v qui fonctionnent sont alors $u = u_0 + u_1$ et $v = v_0 + v_1$.

Exemple 3.36
Écrivons $1000 = u \cdot 7 + v \cdot 5$ avec $u, v \in \mathbb{N}$. D'abord $72 \cdot 7 = 504$ et $100 \cdot 5 = 500$. Nous avons donc

$$1004 = 72 \cdot 7 + 100 \cdot 5. \tag{3.47}$$

Ensuite $4 = 25 - 21 = -3 \cdot 7 + 5 \cdot 5$. Au final,

$$1000 = 75 \cdot 7 + 95 \cdot 5. \tag{3.48}$$

△

Problèmes et choses à faire

Choses à vérifier :
- La proposition 3.43 demande sans doute de parler de la décomposition en facteurs premiers des entiers. Probablement à déplacer.
- Il faut vérifier si le corollaire 3.44 est correct.

Et puis rédiger des démonstrations de tout ce petit monde.

3.5.4 Calcul effectif du PGCD et de Bézout

Source : [15].

Soient a et b, deux entiers que nous allons prendre positifs. Nous allons voir maintenant l'algorithme de **Euclide étendu** qui est capable, pour a et b donnés, de calculer le PGCD de a et b, et un couple de Bézout (u, v) tel que $ua + vb = \text{pgcd}(a, b)$. Ce calcul est indispensable si on veut implémenter RSA (10.4).

Cela se base sur le lemme suivant.

Lemme 3.37.
Soient $a, b \in \mathbb{N}$ et des nombres q et r tels que $a = qb + r$. Alors $\text{pgcd}(a, b) = \text{pgcd}(r, b)$.

Démonstration. Il suffit de voir que les diviseurs communs de a et b sont diviseurs de r et que les diviseurs communs de r et b divisent a.

Si s divise a et b, alors dans l'équation

$$\frac{a}{s} = \frac{qb}{s} + \frac{r}{s}$$

les termes a/s et qb/s sont entiers, donc r/s est aussi entier, et s divise r.

Inversement, si s divise r et b, alors il divise $qb + r$ et donc a. □

Remarque 3.38.
Ce lemme est surtout intéressant lorsque $a = qb + r$ est la division euclidienne de a par b : en effet, dans ce cas $r < b$, et le calcul du PGCD de deux nombres (a et b) est ramené à un calcul de PGCD de deux nombres plus petits (b et r).

L'algorithme pour calculer $\text{pgcd}(a, b)$ consiste à écrire des divisions euclidiennes successives de la manière suivante :

$$a = q_2 b + r_2 \qquad r_2 < b \tag{3.49a}$$
$$b = q_3 r_2 + r_3 \qquad r_3 < r_2 \tag{3.49b}$$
$$\vdots \tag{3.49c}$$

en remarquant que $\text{pgcd}(a, b) = \text{pgcd}(b, r_2) = \text{pgcd}(r_2, r_3)$. Étant donné que les inégalités $r_2 < b$ et $r_3 < r_2$ sont strictes, en continuant ainsi nous finissons sur zéro, c'est à dire qu'il existera un n pour lequel $r_{n+1} = 0$; et donc

$$r_{n-1} = q_{n+1} r_n,$$

et à ce moment nous avons $\text{pgcd}(a, b) = \text{pgcd}(r_{n-1}, r_n) = r_n$.

3.5.4.1 Algorithme d'Euclide pour le PGCD

Écrivons l'algorithme en détail (parce que Bézout, ça va être la même chose en cinq fois plus compliqué). On pose

$$r_0 = a \tag{3.50a}$$
$$r_1 = b \tag{3.50b}$$

(ce qui explique que nous n'ayons pas utilisé r_0 et r_1 précédemment). Ensuite on écrit la division euclidienne $a = q_2 b + r_2$, c'est à dire $r_0 = q_2 r_1 + r_2$. Cela donne r_2 et q_2 en termes de r_0 et r_1 :

$$r_2 = r_0 - q_2 r_1. \tag{3.51}$$

Ensuite, sachant r_2 nous pouvons continuer :

$$r_1 = q_3 r_2 + r_3 \tag{3.52}$$

donne q_3 et $r_3 = r_1 - q_3 r_2$. On continue avec $r_2 = q_4 r_3 + r_4$. Tout cela pour poser la suite

$$\begin{aligned} r_0 &= a \\ r_1 &= b \\ r_k &= q_{k+2} r_{k+1} + r_{k+2} \end{aligned} \tag{3.53}$$

où la troisième définit r_{k+2} et q_{k+2} en fonction de r_k et r_{k+1}, à l'aide du théorème de la division euclidienne. La suite (r_k) ainsi construite est strictement décroissante et à chaque étape le lemme 3.37 et le principe de l'algorithme d'Euclide nous donnent

$$\begin{cases} \operatorname{pgcd}(r_k, r_{k+1}) = \operatorname{pgcd}(r_{k+1}, r_{k+2}) = \operatorname{pgcd}(a, b) & \text{(3.54a)} \\ 0 \leqslant r_{k+1} < r_k. & \text{(3.54b)} \end{cases}$$

La suite étant strictement décroissante, nous prenons r_n, le dernier non nul : $r_{n+1} = 0$. Dans ce cas la dernière équation sera

$$r_{n-1} = q_n r_n \tag{3.55}$$

avec $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(r_n, r_{n-1}) = r_n$.

Exemple 3.39
Calculons le PGCD de 18 et 231. Pour cela nous écrivons les divisions euclidiennes en chaîne :

$$231 = 18 \cdot 12 + 15 \tag{3.56a}$$
$$18 = 1 \cdot 15 + 3 \tag{3.56b}$$
$$15 = 5 \cdot 5 + 0. \tag{3.56c}$$

Donc le PGCD est 3 (le dernier reste non nul). △

3.5.4.2 Algorithme étendu : calcul effectif des coefficients de Bézout

La difficulté est de construire la suite qui donne des coefficients de Bézout. Elle va être construite à l'envers. Nous supposons déjà connaître la liste complète des r_k jusqu'à $r_n = \operatorname{pgcd}(a, b)$, ainsi que la liste complète des divisions euclidiennes

$$r_k = q_{k+2} r_{k+1} + r_{k+2}. \tag{3.57}$$

Nous voulons trouver les couples (u_k, v_k) de telle façon à avoir à chaque étape

$$r_n = u_k r_k + v_k r_{k-1}. \tag{3.58}$$

Notons que c'est à chaque fois r_n que nous construisons. La première équation de type Bézout à résoudre est

$$r_n = u_n r_n + v_n r_{n-1}, \tag{3.59}$$

sachant que $r_{n-1} = q_n r_n$. On pose $v_n = 0$ et $u_n = 1$ et c'est bon. Pour la récurrence, supposons les coefficients u_k et v_k connus, et déterminons les coefficients u_{k-1} et v_{k-1}. Pour ce faire, nous égalons les deux expressions pour r_n :

$$r_n = u_k r_k + v_k r_{k-1} = u_{k-1} r_{k-1} + v_{k-1} r_{k-2}; \tag{3.60}$$

dans laquelle nous substituons $r_{k-2} = q_k - r_{k-1} + r_k$:

$$\begin{aligned} u_k r_k + v_k r_{k-1} &= u_{k-1} r_{k-1} + v_{k-1}(q_k r_{k-1} + r_k) & \text{(3.61)} \\ &= (u_{k-1} + q_k v_{k-1}) r_{k-1} + + v_{k-1} r_k & \text{(3.62)} \end{aligned}$$

et nous égalons les coefficients de r_k et r_{k-1} pour obtenir

$$\begin{cases} v_{k-1} = u_k & \text{(3.63a)} \\ u_{k-1} = v_k - v_{k-1} q_k. & \text{(3.63b)} \end{cases}$$

Dès que u_k et v_k ainsi que q_k sont connus, on peut calculer u_{k-1} et v_{k-1}.

La dernière équation, celle avec $k = 1$, est

$$r_n = u_1 r_1 + v_1 r_0, \tag{3.64}$$

c'est à dire

$$\text{pgcd}(a, b) = u_1 b + v_1 a. \tag{3.65}$$

Nous avons ainsi trouvé des coefficients de Bézout pour a et b.

3.5.5 Décomposition en facteurs premiers

Lemme 3.40 (Lemme de Gauss)**.**
Soient $a, b, c \in \mathbb{Z}$ tels que a divise bc. Si a est premier avec c, alors a divise b.

Démonstration. Vu que a est premier avec c, nous avons $\text{pgcd}(a, c) = 1$ et le théorème de Bézout 3.32 nous donne donc $s, t \in \mathbb{Z}$ tels que $sa + tc = 1$. En multipliant par b, nous avons $sab + tbc = b$. Mais les deux termes du membre de gauche sont multiples de a parce que a divise bc. Par conséquent b est somme de deux multiples de a et donc est multiple de a. □

Lemme 3.41 (Lemme d'Euclide[16])**.**
Si un nombre premier p divise le produit de deux nombres entiers b et c, alors p divise b ou c.

Démonstration. Vu que p est premier, s'il ne divise pas a c'est que $\text{pgcd}(a, p) = 1$. Dans ce cas le lemme de Gauss 3.40 implique que p divise b. □

Le théorème fondamental de l'arithmétique permet de décomposer des nombres en facteurs premiers.

Théorème 3.42 ([17])**.**
Tout entier strictement positif peut être écrit comme un produit de nombres premiers d'une unique façon, à l'ordre près des facteurs.

Démonstration. Soit n un entier positif. Nous prouvons l'existence d'une décomposition en facteurs premiers par récurrence. Le nombre $n = 1$ est le produit d'une famille finie de nombres premiers : la famille vide.

Supposons que tout entier strictement inférieur à un certain entier $n > 1$ est produit de nombres premiers. Deux possibilités apparaissent pour n : il est premier ou non. Si n est premier, et donc produit d'un unique entier premier, à savoir lui-même, le résultat est vrai. Si n n'est pas premier, il se décompose sous la forme kl avec k et l strictement inférieurs à n. Dans ce cas, l'hypothèse de récurrence implique que les entiers k et l peuvent s'écrire comme produits de nombres premiers. Leur produit aussi, ce qui fournit une décomposition de n en produit de nombres premiers. Par application du principe de récurrence, tous les entiers naturels peuvent s'écrire comme produit de nombres premiers.

Nous prouvons maintenant l'unicité. Prenons deux produits de nombres premiers qui sont égaux. Prenons n'importe quel nombre premier p du premier produit. Il divise le premier produit, et, de là, aussi le second. Par le lemme d'Euclide 3.41, p doit alors diviser au moins un facteur dans le second produit. Mais les facteurs sont tous des nombres premiers eux-mêmes, donc p doit être égal à un des facteurs du second produit. Nous pouvons donc simplifier par p les deux produits. En continuant de cette manière, nous voyons que les facteurs premiers des deux produits coïncident précisément. □

En d'autres termes, pour tout entier $n > 1$, il existe une suite finie unique $(p_1, k_1), \ldots (p_r, k_r)$ telle que :

(1) les p_i sont des nombres premiers tels que, si $i < j$, alors $p_i < p_j$;

(2) les k_i sont des entiers naturels non nuls ;

(3) $n = \prod_{i=1}^{r} p_i^{k_i}$.

Proposition 3.43.
Soient $a, b \in \mathbb{Z}^*$. Si

$$a = \epsilon \prod_{p \text{ premiers}} p^{\mu(p)} \quad et \quad b = \epsilon' \prod_{p \text{ premier}} p^{\nu(p)}, \tag{3.66}$$

alors

$$\operatorname{pgcd}(a,b) = \prod p^{\min\{\mu(p),\nu(p)\}} \tag{3.67a}$$

$$\operatorname{ppcm}(a,b) = \prod p^{\max\{\mu(p),\nu(p)\}} \tag{3.67b}$$

Corollaire 3.44 ([1])**.**
Un élément $m \in \mathbb{Z}^*$ vérifie $m \leqslant p^n$ et $\operatorname{pgcd}(m, p^n) \neq 1$ si et seulement si $m = qp$ pour un certain $q \leqslant p^{n-1}$.

Lemme 3.45.
Un entier $n \geqslant 1$ se décompose de façon unique en produit de la forme $n = qm^2$ où q est un entier sans facteurs carrés et m, un entier.

Démonstration. Pour $n = 1$, c'est évident. Nous supposons $n \geqslant 2$.

En ce qui concerne l'existence, nous décomposons n en facteurs premiers [5] et nous séparons les puissances paires des puissances impaires :

$$n = \prod_{i=1}^{r} p_p^{2\alpha_i} \prod_{j=1}^{s} q_j^{2\beta_j+1} \tag{3.68a}$$

$$= \underbrace{\left(\prod_{i=1}^{r} p_i^{2\alpha_i} \prod_{j=1}^{s} q^{2\beta_j} \right)}_{m^2} \underbrace{\prod_{j=1}^{s} q_j}_{q} . \tag{3.68b}$$

Nous passons à l'unicité. Supposons que $n = q_1 m_1^2 = q_2 m_2^2$ avec q_1 et q_2 sans facteurs carrés (dans leur décomposition en facteurs premiers). Soit $d = \operatorname{pgcd}(m_1, m_2)$ et k_1, k_2 définis par $m_1 = dk_1$, $m_2 = dk_2$. Par construction, $\operatorname{pgcd}(k_1, k_2) = 1$. Étant donné que

$$n = q_1 d^2 k_1^2 = q_2 d^2 k_2^2, \tag{3.69}$$

nous avons $q_1 k_1^2 = q_2 k_2^2$ et donc k_1^2 divise $q_2 k_2^2$. Mais k_1 et k_2 n'ont pas de facteurs premiers en commun, donc k_1^2 divise q_2, ce qui n'est possible que si $k_1 = 1$ (parce que k_1^2 n'a que des facteurs premiers alors que q_2 n'en a pas). Dans ce cas, $d = m_1$ et m_1 divise m_2. Si $m_2 = l m_1$ alors l'équation (3.69) se réduit à $n = q_1 m_1^2 = q_2 l^2 m_1^2$ et donc

$$q_1 = q_2 l^2, \tag{3.70}$$

ce qui signifie $l = 1$ et donc $m_1 = m_2$. □

Dans $\mathbb{N}$, il y a assez bien de nombres premiers. Nous allons voir maintenant que la somme des inverses des nombres premiers diverge. Pour comparaison, la somme des inverses des carrés converge. Il y a donc plus de nombres premiers que de carrés.

3.5.6 Écriture des fractions

Théorème 3.46.
Tout élément de $\mathbb{Q}^+$ s'écrit de façon unique comme quotient de deux entiers premiers entre eux.

Démonstration. En deux parties

5. Théorème 3.42.

Unicité Supposons avoir $\frac{a}{b} = \frac{c}{d}$ avec $\text{pgcd}(a,b) = \text{pgcd}(c,d) = 1$. Nous avons

$$ad = bc \tag{3.71}$$

donc

(1) a divise bc mais est premier avec b donc a divise c par le lemme de Gauss 3.40.

(2) c divise ad mais est premier avec d donc c divise a par le lemme de Gauss 3.40.

En conclusion a divise c et c divise a, ergo $a = c$. L'égalité $b = d$ est alors immédiate.

Existence Soit le quotient $\frac{a}{b}$. Nous avons

$$\frac{a}{b} = \frac{a/\text{pgcd}(a,b)}{b/\text{pgcd}(a,b)}, \tag{3.72}$$

qui est encore un quotient d'entiers parce que $\text{pgcd}(a,b)$ divise aussi bien a que b. Il faut montrer que les nombres $a/\text{pgcd}(a,b)$ et $b/\text{pgcd}(a,b)$ sont premiers entre eux. Pour cela nous supposons que k est un diviseur commun. En particulier, les nombres $a/k\,\text{pgcd}(a,b)$ et $b/k\,\text{pgcd}(a,b)$ sont des entiers, ce qui fait que $k\,\text{pgcd}(a,b)$ est un diviseur commun de a et b. Étant donné que $\text{pgcd}(a,b)$ est le plus grand tel diviseur, nous devons avoir $k\,\text{pgcd}(a,b) = \text{pgcd}(a,b)$ c'est à dire que $k = 1$. Donc les nombres $a/\text{pgcd}(a,b)$ et $b/\text{pgcd}(a,b)$ sont premiers entre eux.

□

Proposition 3.47.
Les entiers p et q sont premiers entre eux si et seulement si p^2 et q^2 sont premiers entre eux.

Démonstration. Si p^2 et q^2 sont premiers entre eux, par le théorème de Bézout 3.32 il existe $a, b \in \mathbb{Z}$ tels que

$$ap^2 + bq^2 = 1 \tag{3.73}$$

Dans ce cas, $(ap)p + (bq)q = 1$, ce qui montre (par encore Bézout, mais l'autre sens) que p et q sont premiers entre eux.

Réciproquement, supposons que p et q ne sont pas premiers entre eux. Alors $\text{pgcd}(p,q) = k \neq 1$. L'entier k divise p et donc p^2 ; et l'entier k divise q et donc q^2. Au final, k divise p^2 et q^2, ce qui montre que p^2 et q^2 ne sont pas premiers entre eux. □

Une des conséquences de ces résultats sera le fait que $\sqrt{n}$ est irrationnelle dès que n n'est pas un carré parfait, théorème 12.24.

3.5.7 Équation diophantienne linéaire à deux inconnues

Soient a, b et c des entiers naturels donnés. Nous considérons l'équation

$$ax + by = c \tag{3.74}$$

à résoudre[18] pour $(x,y) \in \mathbb{N}^2$.

Si a ou b est nul, c'est facile ; nous supposons donc que a et b sont tout deux non nuls. Nous commençons par simplifier l'équation en cherchant les diviseurs communs. Soit $d = \text{pgcd}(a,b)$ et notons $a = da'$, $b = db'$. Nous avons déjà l'équation

$$d(a'x + b'y) = c, \tag{3.75}$$

et donc si c n'est pas un multiple de d, il n'y a pas de solutions[6]. Si par contre c est un multiple de d alors nous écrivons $c = c'd$ et l'équation devient

$$a'x + b'y = c' \tag{3.76}$$

6. Exemple : $8x + 2y = 9$. Le membre de gauche est certainement un nombre pair et il n'y a donc pas de solutions.

C'est maintenant que nous utilisons le théorème de Bézout 3.32 : vu que a' et b' sont premiers entre eux, nous avons la relation $a'u + b'v = 1$ pour certains[7] nombres entiers u et v. Nous récrivons notre équation sous la forme $a'x + b'y = c'(a'u + b'v)$ et rassemblons les termes :

$$a'(x - c'u) = b'(c'v - y). \tag{3.77}$$

C'est à dire que si (x, y) est une solution, alors a' divise $b'(c'v - y)$. Mais comme a' et b' sont premiers entre eux, le nombre a' doit forcément[8] diviser $c'v - y$. Disons $c'v - y = ka'$. Alors $a'(x - c'u) = b'ka'$ et donc

$$x = b'k + c'u. \tag{3.78}$$

Nous trouvons alors une expression pour y en injectant cela dans $a'x + b'y = c'$ et en utilisant Bézout : $a'c'u = (1 - b'v)c'$. Au final nous avons prouvé que toutes les solutions sont de la forme

$$\begin{cases} x = b'k + c'u & \text{(3.79a)} \\ y = vc' - a'k & \text{(3.79b)} \end{cases}$$

avec $k \in \mathbb{Z}$. Si nous voulons réellement seulement des solutions dans $\mathbb{N}$ et non dans $\mathbb{Z}$, il faut seulement un peu restreindre les valeurs de k. Il en reste un nombre fini parce que l'équation pour x borne k vers le bas tandis que celle pour y borne k vers le haut.

Par ailleurs, il est très vite vérifié que tous les couples (x, y) de la forme (3.79) sont solutions.

Exemple 3.48
Résoudre l'équation $2x + 6y = 52$.

Nous pouvons factoriser 2 dans le membre de gauche et simplifier alors toute l'équation par 2 :

$$x + 3y = 26. \tag{3.80}$$

Nous cherchons une relation de Bézout pour $u+3v = 1$. Ce n'est heureusement pas très compliqué : $u = -5$, $v = 2$. Nous pouvons alors écrire

$$x + 3y = 26 \times (-5 + 3 \times 2), \tag{3.81}$$

et donc $x + 5 \times 26 = 3(y - 26 \times 6)$, et en posant $k = y - 26 \times 6$ nous avons

$$x = 3k - 130. \tag{3.82}$$

En injectant nous trouvons l'équation $3k - 130 + 3y = 26$ et donc

$$y = 52 - k. \tag{3.83}$$

△

3.5.8 Quotients

Nous écrivons $a = b \mod p$ essentiellement s'il existe $k \in \mathbb{Z}$ tel que $b + kp = a$. Plus généralement nous notons $[a]_p = \{a + kp | k \in \mathbb{Z}\}$ et l'écriture «$a = n \mod p$» peut tout autant signifier $a = [b]_p$ que $a \in [b]_p$. La différence entre les deux est subtile mais peut de temps en temps valoir son pesant d'or.

Proposition 3.49.
Soit $n \in \mathbb{N}$. Le groupe $\mathbb{Z}/n\mathbb{Z}$ est monogène. Si $n \neq 0$, le groupe $\mathbb{Z}/n\mathbb{Z}$ est cyclique d'ordre n.

7. Nous avons décrit un algorithme pour les trouver en démontrant l'équation 3.65.
8. C'est Gauss 3.40 qui le dit, et vous savez que lorsque Gauss dit, c'est *forcément* vrai.

Démonstration. Nous considérons la surjection canonique $\mu\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$. Si $a \in \mathbb{Z}$, alors $\mu(a) = a\mu(1)$. Par conséquent $\mathbb{Z}/n\mathbb{Z} = \text{gr}\big(\mu(1)\big)$ parce que tout groupe contenant $\mu(1)$ contient tous les multiples de $\mu(1)$, et par conséquent contient $\mu(\mathbb{Z}) = \mathbb{Z}/n\mathbb{Z}$.

Soit $x \in \mathbb{Z}/n\mathbb{Z}$ et considérons x_0, le plus petit naturel représentant x. Nous notons $x = [x_0]$. Le théorème de la division euclidienne 3.25 donne l'existence de q et r avec $0 \leqslant r < n$ et $q \geqslant 0$ tels que

$$x_0 = nq + r. \tag{3.84}$$

Nous avons $[x_0] = [r] = \mu(r)$ parce que $x_0 - r$ est un multiple de n. Nous avons donc $[x_0] \in \mu(\mathbb{N}_{n-1})$. Par conséquent

$$\mathbb{Z}/n\mathbb{Z} = \mu(\mathbb{Z}) = \mu(\mathbb{N}_{n-1}). \tag{3.85}$$

La restriction $\mu\colon \mathbb{N}_{n-1} \to \mathbb{Z}/n\mathbb{Z}$ est donc surjective. Montrons qu'elle est également injective. Si $\mu(x_0) = \mu(x_1)$, alors $x_1 = x_0 + kn$. Si nous supposons que $x_1 > x_0$, alors $k > 0$ et si $x_0 \in \mathbb{N}_{n-1}$, alors $x_1 > n - 1$.

L'ordre de $\mathbb{Z}/n\mathbb{Z}$ est donc le même que le cardinal de $\mathbb{N}_{n-1}$, c'est à dire n. Le groupe $\mathbb{Z}/n\mathbb{Z}$ est donc fini, d'ordre n et monogène parce que $\mathbb{Z}/n\mathbb{Z} = \text{gr}(\mu(1))$. Il est donc cyclique. □

Lemme 3.50 ([2])**.**
Soit $q \in \mathbb{N}$ avec $q \geqslant 2$. Soient $n, d \in \mathbb{N}$ tels que $q^d - 1 \mid q^n - 1$. Alors $d \mid n$.

Démonstration. Par le théorème de division euclidienne 3.25, il existe $a, b \in \mathbb{Z}$ tels que $n = ad + b$ avec $0 \leqslant b < d$. En remarquant que $q^d \in [1]_{q^d-1}$ nous avons

$$q^n = (q^d)^a q^b \in [1]_{q^d-1} q^b = [q^b]_{q^d-1}. \tag{3.86}$$

Pour cela nous avons utilisé d'abord le fait que si $a \in [z]_k$, alors $a^n \in [z^n]_k$ et ensuite le fait que $[1]_k x = [x]_k$. D'autre part l'hypothèse $q^d - 1 \mid q^n - 1$ implique

$$q^n \in [1]_{q^d-1}. \tag{3.87}$$

Par conséquent le nombre q^n est à la fois dans $[q^b]_{q^d-1}$ et dans $[1]_{q^d-1}$. Cela implique que

$$[1]_{q^d-1} = [q^b]_{q^d-1}, \tag{3.88}$$

ou encore que $q^b \in [1]_{q^d-1}$ ou encore que $q^d - 1 \mid q^b - 1$.

Étant donné que $b < d$ et que $q \geqslant 2$, nous avons que $q^b - 1 < q^d - 1$; donc pour que $q^d - 1$ divise $q^b - 1$, il faut que $q^b - 1$ soit zéro, c'est à dire $b = 0$.

Mais dire $b = 0$ revient à dire que $d \mid n$, ce qu'il fallait démontrer. □

3.6 Indice d'un sous-groupe et ordre des éléments

Lemme 3.51.
Lorsque H est normal dans G, alors la définition

$$[a] \cdot [b] = [ab] \tag{3.89}$$

définit une loi de groupe sur l'ensemble G/H.

Démonstration. Le neutre est $[e]$ et l'associativité ne pose pas plus de problèmes que l'existence d'un inverse. Le point à vérifier est que la formule (3.89) est une bonne définition : $[ah] \cdot [bh'] = [ab]$ pour tout $h, h' \in H$. Nous avons :

$$[ah] \cdot [ah'] = [ahah'] = [ahb]. \tag{3.90}$$

Pour montrer que cela et $[ab]$, l'astuce est d'introduire bb^{-1} à côté du a :

$$[ahb] = [abb^{-1}hb] = [ab] \tag{3.91}$$

parce que $b^{-1}hb \in H$ du fait que H soit normal dans G. □

Exemple 3.52([19])
Il ne faudrait pas croire que le groupe quotient G/H est forcément un sous-groupe de G. Par exemple le quotient $\mathbb{Z}/2\mathbb{Z}$ est l'ensemble $\{0,1\}$ muni de l'addition. En particulier $1+1=0$, ce qui est évidemment faux dans $\mathbb{Z}$. Le groupe $(\mathbb{Z},+)$ ne possède aucun élément d'ordre 2.

Il n'en est pas moins vrai que l'application

$$\begin{aligned} f\colon G &\to G/H \\ g &\mapsto [g] \end{aligned} \tag{3.92}$$

est un morphisme de groupes. △

Définition 3.53.
*Si H est un sous-groupe d'un groupe fini l'**indice** de H dans G est le nombre $|G|/|H|$, souvent noté $|G:H|$.*

Le théorème de Lagrange dira en particulier que l'indice est toujours un nombre entier. C'est à ne pas confondre avec le degré d'une extension de corps (définition 5.53).

Théorème 3.54 (Théorème de Lagrange)**.**
Soit H un sous-groupe du groupe fini G. Alors

(1) L'ordre de H divise l'ordre de G.

(2) Les trois nombres suivants sont égaux :
- *le nombre de classes de H à gauche,*
- *le nombre de classes de H à droite,*
- *l'indice de H dans G.*

En particulier si H est distingué dans G nous avons

$$|G/H| = \frac{|G|}{|H|}. \tag{3.93}$$

Démonstration. Nous commençons par montrer que les classes de H ont toutes les même nombre d'éléments que H. En effet pour chaque $g \in G$ nous avons la bijection

$$\begin{aligned} \varphi\colon H &\to gH \\ h &\mapsto gh. \end{aligned} \tag{3.94}$$

L'injectivité de φ est le fait que $gh = gh'$ implique $h = h'$. La surjectivité est par définition de la classe.

Les classes à gauche formant une partition de G, le cardinal de G est le produit de la taille des classes par le nombre de classes :

$$|G| = |H| \cdot \text{ nombre de classes.} \tag{3.95}$$

En particulier nous voyons que $|H|$ divise $|G|$.

La dernière formule exprime simplement que G/H est par définition le nombre de classes de H à gauche (ou à droite) dans G. □

Corollaire 3.55.
L'ordre d'un élément d'un groupe fini divise l'ordre du groupe. En particulier dans un groupe d'ordre n tous les éléments vérifient $g^n = e$.

Démonstration. Soit G un groupe fini et considérons, à $g \in G$ fixé, le sous-groupe

$$H = \{g^k \text{ tel que } k \in \mathbb{N}\}. \tag{3.96}$$

Par le théorème de Lagrange 3.54, l'ordre de H divise $|G|$, mais l'ordre de H est le plus petit k tel que $g^k = e$, c'est à dire l'ordre de g. □

Théorème 3.56 (Théorème de Cauchy[20])**.**
Soit un groupe fini d'ordre n. Pour tout diviseur premier p de n, le groupe G possède au moins un élément d'ordre p.

Le lemme suivant indique que sous hypothèse de commutativité, l'ordre d'un élément est une notion multiplicative.

Lemme 3.57 ([21])**.**
Soit G un groupe et $a, b \in G$ tels que $ab = ba$ d'ordres respectivement r et s, deux nombres premiers entre eux. Alors l'élément ab est d'ordre rs.

Démonstration. Étant donné que $(ab)^{rs} = a^{rs}b^{rs} = 1$, l'ordre de ab divise rs. Et vu que r et s sont premiers entre eux, l'ordre de ab s'écrit sous la forme r_1s_1 avec $r_1 \mid r$ et $s_1 \mid s$. Nous avons

$$a^{r_1s_1}b^{r_1s_1} = (ab)^{r_1s_1} = 1, \tag{3.97}$$

que nous élevons à la puissance r_2 où r_2 est définit en posant $r = r_1r_2$:

$$a^{rs_1}b^{rs_1} = 1. \tag{3.98}$$

Et comme $a^{rs_1} = 1$, nous concluons que $b^{rs_1} = 1$. Donc $s \mid rs_1$. Par le lemme de Gauss (3.40), nous avons en fait $s \mid s_1$. Vu qu'on a aussi $s_1 \mid s$, nous avons $s = s_1$.

Le même type d'argument donne $r = r_1$, et finalement l'ordre de ab est $r_1s_1 = rs$. □

Lemme 3.58.
L'ensemble des ordres d'un groupe commutatif *est stable par PPCM.*

Autrement dit, si $x \in G$ est d'ordre r et si $y \in G$ est d'ordre s, alors il existe un élément d'ordre $\text{ppcm}(r, s)$.

Démonstration. Soit $m = \text{ppcm}(r, s)$. Afin d'écrire m sous une forme pratique, nous considérons les décompositions en facteurs premiers de r et s :

$$r = \prod_{i=1}^{k} p_i^{\alpha_i} \tag{3.99a}$$

$$s = \prod_{i=1}^{k} p_i^{\beta_i} \tag{3.99b}$$

où $\{p_i\}_{i=1\ldots k}$ est l'ensemble des nombres premiers arrivant dans les décompositions de r et de s. Si nous posons

$$r' = \prod_{\substack{i=1 \\ \alpha_1 > \beta_i}}^{k} p_i^{\alpha_i} \tag{3.100a}$$

$$s' = \prod_{\substack{i=1 \\ \alpha_i \leqslant \beta_i}}^{k} p_i^{\beta_i}, \tag{3.100b}$$

alors $\text{ppcm}(r, s) = r's'$ et r' et s' sont premiers entre eux. L'élément $x^{r/r'}$ est d'ordre r' et l'élément $y^{s/s'}$ est d'ordre s'. Maintenant nous utilisons le fait que G soit commutatif et le lemme 3.57 pour conclure que l'ordre de $x^{r/r'}y^{s/s'}$ est $r's' = m$. □

Lemme 3.59 ([22])**.**
Un sous-groupe d'indice 2 est un sous-groupe normal.

Démonstration. Si H est un tel sous-groupe d'un groupe G, alors G possède exactement deux classes à gauche par rapport à H (théorème de Lagrange 3.54) et se partitionne donc en deux parties : $G = H \cup xH$ avec $x \notin H$. De même pour les classes à droite : $G = H \cup Hx$. Puisque la classe à droite Hx n'est pas H, on a $xH = Hx$, et H est normal dans G par la proposition 3.11. □

Lemme 3.60 ([23]).
Soit H, un sous-groupe normal d'indice m de G. Alors pour tout $a \in G$ nous avons $a^m \in H$.

Démonstration. Par définition de l'indice, le groupe G/H est d'ordre m. Donc si $[a] \in G/H$, nous avons $[a]^m = [e]$, ce qui signifie $[a^m] = [e]$, ou encore $a^m \in H$. □

Proposition 3.61 ([23]).
Soit un groupe fini G et H, un sous-groupe normal d'ordre n et d'indice m avec m et n premiers entre eux. Alors H est l'unique sous-groupe de G à être d'ordre n.

Notons que cette proposition ne dit pas qu'il existe un sous-groupe d'ordre n et d'indice m. Il dit juste que s'il y en a un et s'il est normal, alors il n'y en a pas d'autres.

Démonstration. Soit H' un sous-groupe d'ordre n. Si $h \in H'$ alors $h^n = 1$ et $h^m \in H$ (lemme 3.60). Étant donné que m et n sont premiers entre eux, par le théorème de Bézout 3.32, il existe $a, b \in \mathbb{Z}$ tels que

$$am + bn = 1. \tag{3.101}$$

Du coup $h = h^1 = (h^m)^a(h^n)^b$. En tenant compte du fait que $h^n = 1$ et $h^m \in H$, nous avons $h \in H$. Ce que nous venons de prouver est que $H' \subset H$ et donc que $H = H'$ parce que $|H'| = |H| = |G|/m$. □

3.7 Suite de composition

Définition 3.62.
*Soit G un groupe. Une **suite de composition** pour G est une suite finie de sous-groupes $(G_i)_{i=0,\dots,n}$ telle que*

$$\{e\} = G_n \subseteq G_{n-1} \subseteq \dots \subseteq G_1 \subseteq G_0 = G \tag{3.102}$$

*et telle que G_{i+1} est normal[9] dans G_i. Les groupes G_i/G_{i+1} sont les **quotients** de la suite de composition.*

*Une suite de **Jordan-Hölder** est une suite de composition dont tous les quotients sont simples.*

L'objet de nos prochaines pérégrinations mathématiques est de montrer que tout groupe fini admet une suite de Jordan-Hölder (théorème 3.67).

Lemme 3.63 (du papillon[24]).
Soit G un groupe et des sous-groupes A et B. Soit A' normal dans A et B' normal dans B. Alors

(1) $A'(A \cap B')$ est normal dans $A'(A \cap B)$

(2) $(A' \cap B)B'$ est normal dans $(A \cap B)B'$

(3) Nous avons les isomorphismes de groupes

$$\frac{A'(A \cap B)}{A'(A \cap B')} \simeq \frac{(A \cap B)B'}{(A' \cap B)B'} \simeq \frac{B'(A \cap B)}{B'(A' \cap B)}. \tag{3.103}$$

Démonstration. Nous n'allons pas démontrer chacun des points ; nous renvoyons au fameux «preuve très similaire dans les autres cas» pour plus de justifications.

Commençons par montrer que $A'(A \cap B')$ est un groupe. Si $a, b \in A'$ et $x, y \in A \cap B'$,

$$axby = xx^{-1}axbx^{-1}xy \tag{3.104}$$

En utilisant la normalité, $x^{-1}ax \in A'$, donc $xx^{-1}axbx^{-1} \in A'$ et donc le tout est dans $A'(A \cap B')$. L'ensemble $A'(A \cap B')$ est également stable pour l'inverse parce que

$$x^{-1}a^{-1} = \underbrace{x^{-1}a^{-1}x}_{\in A'}x^{-1}. \tag{3.105}$$

9. Nous rappelons au cas où que «normal» signifie «distingué».

Nous montrons maintenant que $A'(A\cap B')$ est normal dans $A'(A\cap B)$. Soient $a,b\in A'$, $x\in A\cap B'$ et $f\in A\cap B$. Alors

$$\begin{aligned}(bf)^{-1}(ax)(bf)&=(bf)^{-1}(a\underbrace{xbx^{-1}}_{=c\in A'}xf) && (3.106a)\\ &=f^{-1}b^{-1}acxf && (3.106b)\\ &=f^{-1}b^{-1}acf\ \underbrace{f^{-1}xf}_{=y\in A\cap B'} && (3.106c)\\ &=\underbrace{f^{-1}b^{-1}acf}_{\in A'}y && (3.106d)\\ &\in A'(A\cap B'). && (3.106e)\end{aligned}$$

Pour prouver l'isomorphisme

$$\frac{A'(A\cap B)}{A'(A\cap B')}=\frac{(A\cap B)B'}{(A'\cap B)B'}, \tag{3.107}$$

nous allons utiliser le deuxième théorème d'isomorphisme (3.24(5)). Que nous appliquons à $H=A\cap B$ et $N=A'(A\cap B')$. La vérification que H normalise N est usuelle. Nous commençons par écrire

$$\frac{A'(A\cap B')(A\cap B)}{A'(A\cap B')}\simeq\frac{A\cap B}{A\cap B\cap A'(A\cap B')}. \tag{3.108}$$

Pour simplifier un peu cette expression nous prouvons d'abord que

$$(A\cap B)\cap A'(A\cap B')=(A'\cap B)(A\cap B'). \tag{3.109}$$

L'inclusion $\supset$ est facile. Pour l'autre sens, étant donné que $A'(A\cap B')\subset A$ nous avons

$$A\cap B\cap A'(A\cap B)=B\cap A'(A\cap B). \tag{3.110}$$

Un élément de $B\cap A'(A\cap B)$ est un élément de B qui s'écrit sous la forme $s=ax$ avec $a\in A'$ et $x\in A\cap B'$. Nous avons alors $a=sx^{-1}$ avec $s\in B$ et $x^{-1}\in B'$. Par conséquent $a\in B$ et donc $a\in A'\cap B$. Nous avons donc

$$(A\cap B)\cap A'(A\cap B')=B\cap A'(A\cap B)\subset(A'\cap B)(A\cap B'), \tag{3.111}$$

et donc l'égalité (3.109). Toujours dans l'idée de simplifier (3.108) nous remarquons que $A\cap B'$ est un sous-ensemble de $A\cap B'$, donc $A'(A\cap B')(A\cap B)=A'(A\cap B)$. Il reste donc

$$\frac{A'(A\cap B)}{A'(A\cap B')}=\frac{A\cap B}{(A'\cap B)(A\cap B')}. \tag{3.112}$$

Étant donné que les hypothèses sur A et B sont symétriques, le membre de droite peut aussi s'écrire en inversant A et B. Nous en sommes à

$$\frac{B'(A\cap B)}{B'(A'\cap B)}=\frac{A'(A\cap B)}{A'(A\cap B')}. \tag{3.113}$$

Nous devons encore justifier $B'(A\cap B)=(A\cap B)B'$ et $B'(A'\cap B)=(A'\cap B)B'$. Faisons le premier et laissons le second au lecteur. Si $b\in B'$ et $x\in A\cap B$, alors

$$bx=x\underbrace{x^{-1}bx}_{\in B'}\in(A\cap B)B'. \tag{3.114}$$

□

Proposition 3.64.
Si G est un groupe fini et que (G_i) est une suite de composition pour G, alors l'ordre de G est le produit des ordres de ses quotients.

Démonstration. Étant donné que G_{i+1} est toujours normal dans G_i, le théorème de Lagrange (3.54) s'applique et nous avons à chaque pas de la suite de composition nous avons

$$|\frac{G_i}{G_{i+1}}| = \frac{|G_i|}{|G_{i+1}|} \tag{3.115}$$

et il suffit d'écrire $|G|$ de façon télescopique :

$$|G| = \prod_{0\leqslant i\leqslant n-1} \frac{|G_i|}{|G_{i+1}|} \tag{3.116}$$

□

Nous disons que les deux suites de composition $(G_i)_{0\leqslant i\leqslant r}$ et $(G_j)_{0\leqslant j\leqslant s}$ sont **équivalentes** si $r = s$ et s'il existe une permutation $\sigma \in S_{r-1}$ telle que

$$\frac{G_i}{G_{i+1}} \simeq \frac{H_{\sigma(i)}}{H_{\sigma(i)+1}}. \tag{3.117}$$

Proposition 3.65 (Schreider)**.**
Deux suites de composition d'un même groupe admettent des raffinements équivalents.

Démonstration. Soient les suites de composition

$$\{e\} = G_m \subseteq \ldots \subseteq G_1 \subseteq G_0 = G \tag{3.118a}$$
$$\{e\} = H_m \subseteq \ldots \subseteq H_1 \subseteq H_0 = G \tag{3.118b}$$

Nous raffinons la suite (G_i) en remplaçant $G_{i+1} \subseteq G_i$ par

$$G_{i+1} = G_{i+1}(G_i \cap H_n) \subset G_{i+1}(G_i \cap H_{n-1}) \subseteq \ldots \subseteq G_{i+1}(G_i \cap H_0) = G_i, \tag{3.119}$$

et de même pour (H_j). Le groupe $G_{i+1}(G_i \cap H_k)$ est normal dans $G_{i+1}(G_i \cap H_{k-1})$ parce que G_{i+1} étant normal dans G_i et H_k dans H_{k-1}, le lemme 3.63 s'applique. Nous avons donc bien défini un raffinement.

Nous devons maintenant prouver que les deux raffinements ainsi construits sont des suites de composition équivalentes. D'abord elles ont la même longueur mn parce que chacun des m éléments de la suite (G_i) a été remplacé par n éléments et inversement, chacun de n éléments de la suite (H_j) a été remplacé par m éléments.

Par ailleurs, les quotients du raffinement de (G_i) sont de la forme

$$\frac{G_{i+1}(G_i \cap H_k)}{G_{i+1}(G_i \cap H_{k+1})} \simeq \frac{H_{k+1}(H_k \cap G_i)}{H_{k+1}(H_k \cap G_{i+1})} \tag{3.120}$$

en vertu du lemme du papillon (3.63). Le membre de droite de (3.120) est un des quotients du raffinement de (H_j). □

Lemme 3.66 (Schreider strictement décroissant)**.**
Soient Σ_1 et Σ_2, deux suites de composition strictement décroissantes du groupe G. Alors elles admettent des raffinements équivalents strictement décroissants.

Démonstration. Par hypothèse, Σ_1 et Σ_2 n'ont pas de répétitions. Soient Σ_1'' et Σ_2'', des raffinements équivalents donnés par le lemme de Schreider. Étant donné que ce sont des suites de composition équivalentes, elles ont le même nombre de quotients réduits à $\{e\}$, c'est à dire le même nombre de répétitions.

Les suites Σ_1' et Σ_2' obtenues en retirant les répétitions de Σ_1'' et Σ_2'' sont des raffinements équivalents de Σ_1 et Σ_2 et strictement décroissants. □

Théorème 3.67 (Jordan-Hölder)**.**
Tout groupe fini admet une suite de Jordan-Hölder.

Deux suites de Jordan-Hölder sont équivalentes.

Démonstration. Nous ne prouvons que le second point.

Par définition, une suite de Jordan-Hölder n'a pas de raffinement strictement décroissant (à part elle-même) parce que G_{i+1} est normal maximum dans G_i. Si Σ_1 et Σ_2 sont des suites de Jordan-Hölder nous pouvons considérer les raffinements équivalents strictement décroissants Σ_1' et Σ_2' du lemme de Schreider 3.66. Nous avons $\Sigma_1' \sim \Sigma_2'$, mais par ce que nous venons de dire à propos de la maximalité, $\Sigma_1' = \Sigma_1$ et $\Sigma_2' = \Sigma_2$. D'où le résultat. □

3.8 Groupes résolubles

Définition 3.68.
*Le groupe G est **résoluble** s'il existe une suite finie de sous-groupes G_i*

$$\{e\} = G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0 = G \tag{3.121}$$

avec G_i normal dans G_{i+1} et G_i/G_{i+1} abélien.

Il s'agit d'un groupe qui admet une suite de composition [10] dont les quotients sont abéliens.

Lemme 3.69.
Soit G un groupe et H un sous-groupe normal. Le groupe G/H est abélien si et seulement si $D(G) \subset H$.

Démonstration. Les proposition suivantes sont équivalentes :
- Le groupe G/H est abélien
- pour tout $x, y \in G$, $\bar{x}\bar{y} = \bar{y}\bar{x}$
- $\bar{x}\bar{y}\bar{x}^{-1}\bar{y}^{-1} = \bar{e}$
- $\overline{xyx^{-1}y^{-1}} = \bar{e}$
- $[x, y] \in H$
- $D(G) \subset H$.

□

Proposition 3.70.
Un groupe est résoluble si et seulement si sa suite dérivée termine sur $\{e\}$.

Démonstration. Grâce au lemme 3.20 et à la proposition 3.21, si la suite dérivée termine sur $\{e\}$ alors la suite dérivé est une suite qui répond aux conditions de la définition 3.68 de groupe résoluble.

Il faut donc encore montrer le sens direct. Nous supposons que G est un groupe résoluble et nous étudions sa suite dérivée. Nous avons une suite

$$\{e\} = G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0 = G \tag{3.122}$$

avec G_i/G_{i+1} abélien et G_{i+1} normal dans G_i. Nous allons prouver par récurrence que $D^i(G) \subset G_i$.

Pour $i = 0$ nous avons bien $G \subset G_0$. Notre hypothèse de récurrence est :

$$D^i(G) \subset G_i \tag{3.123}$$

Par le lemme 3.69 nous avons aussi

$$D(G_i) \subset G_{i+1}. \tag{3.124}$$

En dérivant (3.123) et en tenant compte de (3.124), $D^{i+1}(G) \subset D(G_i) \subset G_{i+1}$. Donc par récurrence nous avons bien $D^k(G) \subset G_k$ pour tout k. Mais $G_r = \{e\}$ pour un certain r, donc pour ce r nous avons $D^r(G) = \{e\}$, ce qu'il fallait. □

Proposition 3.71.
Soient des groupes G et H. Nous supposons que G est résoluble et nous considérons un homomorphisme $\psi\colon G \to H$. Alors $\psi(G)$ est résoluble.

10. Voir définition 3.62.

Démonstration. Vu que G est résoluble, il existe une suite de sous-groupes G_i tels que

$$\{e\} = G_n \subset G_{n-1} \subset \ldots \subset G_1 \subset G_0 = G \tag{3.125}$$

avec G_i normal dans G_{i+1} et G_i/G_{i+1} abélien. Nous posons $\psi(G)_i = \psi(G_i)$ et nous avons $\psi(G)_n = \psi\big(\{e\}\big) = \{e\}$ ainsi que $\psi(G)_0 = \psi(G)$; donc

$$\{e\} = \psi(G)_n \subset \psi(G)_{n-1} \subset \ldots \subset \psi(G)_1 \subset \psi(G)_0 = \psi(G). \tag{3.126}$$

Les faits que $\psi(G)_i$ soit normal dans $\psi(G)_{i+1}$ et que $\psi(G)_i/\psi(G)_{i+1}$ soit abélien est directement la proposition 3.16.

□

3.9 Action de groupes

Définition 3.72 (Thème 57)**.**
Une ***action de groupe*** *G sur un ensemble E est la donnée, pour chaque élément $g \in G$, d'une fonction $\phi_g : E \to E$, de telle sorte que :*

$$\begin{aligned} \phi_e(x) &= x, & \forall x \in E;\\ \phi_{gh}(x) &= \phi_g(\phi_h(x)), & \forall g, h \in G, \forall x \in E. \end{aligned}$$

On dit dans ce cas que G ***agit*** *sur E.*

Lemme 3.73.
Pour tout $g \in G$,

(1) L'application $\phi_g \colon E \to E$ est injective,

(2) Pour l'inverse : $\phi_{g^{-1}} = \phi_{g^{-1}}$.

Démonstration. Si $x, y \in E$ sont tels que $\phi_g(x) = \phi_g(y)$ alors en appliquant $\phi_{g^{-1}}$ aux deux membre nous trouvons

$$(\phi_{g^{-1}}\phi_g)(x) = (\phi_{g^{-1}}\phi_g)(y), \tag{3.127}$$

ce qui donne $x = y$ parce que $\phi_{g^{-1}}\phi_g = \phi_{g^{-1}g} = \phi_e = \mathrm{Id}$.

Les dernières trois égalités écrites disent que $\phi_{g^{-1}}$ est l'inversa[11] de ϕ_g. □

Pour alléger les notations, on convient d'écrire $g \cdot x$, voire plus simplement gx au lieu de $\phi_g(x)$. Le deuxième axiome d'action de groupe dit que la notation ghx ne souffre d'aucune ambiguïté.

Définition 3.74 (Quelques notions autour de l'action)**.**
*Si G agit sur un ensemble E, nous notons $G \cdot x$ l'****orbite*** *de $x \in E$ sous l'action de G :*

$$G \cdot x = \{gx \text{ tel que } g \in G\}.$$

Nous notons G_x ou $\mathrm{Fix}(x)$ ou $\mathrm{Fix}(x)$ le ***stabilisateur*** *de x :*

$$Fix(x) = G_x = \mathrm{Fix}(x) = \{g \in G \text{ tel que } g \cdot x = x\}. \tag{3.128}$$

Pour $g \in G$, nous notons enfin $\mathrm{Fix}(g)$ le ***fixateur*** *de g :*

$$\mathrm{Fix}(g) = \{x \in E \text{ tel que } g \cdot x = x\}. \tag{3.129}$$

Définition 3.75.
L'action de G sur E est ***fidèle*** *si l'identité est le seul élément de G à fixer tous les points de E, c'est à dire si $gx = x\, \forall x \in E \Rightarrow g = e$.*

11. Si vous décidez de dire ça a un jury dans un concours, soyez prêts à préciser les domaines.

Un exemple d'action fidèle tout à fait non trivial sera donné avec l'action du groupe modulaire sur le plan de Poincaré dans le théorème 21.93.

Le groupe G agit toujours sur lui même à gauche et à droite. L'action à gauche est $g \cdot h = gh$; celle à droite est $g \cdot h = hg^{-1}$.

Définition 3.76.
L'action ***adjointe*** *définie par $g \cdot h = ghg^{-1}$ est une manière pour un groupe d'agir sur lui-même par automorphismes. Cela est souvent noté $\mathbf{Ad}(g)h = ghg^{-1}$.*

En effet pour tout $g \in G$, l'application $\mathbf{Ad}(g) \colon G \to G$ est un automorphisme de G.

Si H est un sous-groupe de G, nous notons G/H le quotient de G par la relation $g \sim gh$ pour tout $h \in H$. Lorsque la distinction est importante, nous noterons $(G/H)_g$ pour les classes à gauche et $(G/H)_d$ pour les classes à droite.

Nous avons une relation d'équivalence à gauche et une à droite. D'abord

$$x \sim_g y \Leftrightarrow xh = y \tag{3.130}$$

pour un certain $h \in H$. Ensuite

$$x \sim_d y \Leftrightarrow hx = y \tag{3.131}$$

pour un certain $h \in H$.

Le lemme suivant est une généralisation du théorème de Lagrange 3.54.

Lemme 3.77.
L'ensemble $(G/H)_g$ est fini si et seulement si l'ensemble $(G/H)_d$ est fini. S'il en est ainsi, alors $(G/H)_g$ et $(G/H)_d$ ont même cardinal qui vaut l'indice de H dans G.

Démonstration. L'application

$$\begin{aligned} f \colon (G/H)_g &\to (G/H)_d \\ [x]_g &\mapsto [x^{-1}]_d \end{aligned} \tag{3.132}$$

est une bijection bien définie. En effet si $x \sim_g y$, nous avons $h \in H$ tel que $y^{-1}h = x^{-1}$, c'est à dire que $x^{-1} \sim_d y^{-1}$ et f est bien définie. Le fait que f soit surjective est évident. Pour l'injectivité, soient $x, y \in G$ tels que

$$f([x]_g) = f([y]_g). \tag{3.133}$$

Alors $x^{-1} \sim_d y^{-1}$, ce qui implique l'existence de $h \in H$ tel que $hx^{-1} = y^{-1}$, ou encore que $xh^{-1} = y$, ce qui signifie que $x \sim_g y$.

Pour l'énoncé à propos de l'indice, nous procédons en plusieurs étapes simples.

(1) Les classes (les éléments de $(G/H)_g$) forment une partition de G.

(2) Toutes les classes ont le même nombre d'éléments par la bijection

$$\begin{aligned} f \colon [x]_g &\to [y]_g \\ xh &\mapsto yh. \end{aligned} \tag{3.134}$$

(3) Le nombre d'éléments dans une classe est égal à $|H|$ par la bijection

$$\begin{aligned} g \colon [x]_g &\to H \\ xh &\mapsto h. \end{aligned} \tag{3.135}$$

Par conséquent

$$|G| = |H| \cdot \text{nombre de classes} = |H| \cdot \text{cardinal de } (G/H)_g, \tag{3.136}$$

et nous avons bien

$$\text{cardinal de } (G/H)_g = \frac{|G|}{|H|} = |G : H|. \tag{3.137}$$

□

Proposition 3.78 (Orbite-stabilisateur[22])**.**
Soit G un groupe agissant sur un ensemble E et $x \in E$.

(1) Les ensembles $G \cdot x$ et G/G_x sont équipotents.

(2) L'orbite de $\mathrm{Fix}(x)$ est finie si et seulement si $\mathrm{Fix}(x)$ est d'indice fini dans G. Dans ce cas nous avons

$$\mathrm{Card}(G \cdot x) = |G : Fix(x)|. \tag{3.138}$$

Une autre façon d'écrire la même formule :

$$|G| = |\mathrm{Fix}(x)||\mathcal{O}_x|. \tag{3.139}$$

C'est la formule (3.138) qui est nommée formule des classes sur wikipédia.

Démonstration. (1) Soit l'application

$$\begin{aligned} \psi\colon G \cdot x &\to G/G_x \\ a \cdot x &\mapsto [a]. \end{aligned} \tag{3.140}$$

Cette application est bien définie parce que si $a \cdot x = b \cdot x$, alors il existe $h \in G_x$ tel que $b = ah$, et par conséquent $[a] = [b]$. Cette application est une bijection et par conséquent $G \cdot x$ est équipotent à G/G_x.

(2) Soit $y \in \mathcal{O}_x$ et $A_y = \{g \in G \text{ tel que } g \cdot x = y\}$. L'ensemble A_y est une classe à gauche de $\mathrm{Fix}(x)$, par conséquent $|A_y| = |\mathrm{Fix}(x)|$ pour tout $y \in \mathcal{O}_x$. Les A_y pour différents y sont disjoints et nous avons de plus

$$\bigcup_{y \in \mathcal{O}_x} A_y = G. \tag{3.141}$$

Les ensemble A_y divisent donc G en $|\mathcal{O}_x|$ paquets de $|\mathrm{Fix}(x)|$ éléments. D'où la formule (3.139).

□

Corollaire 3.79.
Soit C_g la classe de conjugaison d'un élément g du groupe fini G. Alors

$$\mathrm{Card}(C_g) = |G : Z_G(g)| \tag{3.142}$$

où $Z_G(g)$ est le centralisateur de g dans G[12] *de G.*

Démonstration. Cela est une application de la proposition 3.78 (formule (3.138)) dans le cas de l'action adjointe de G sur lui-même.

En effet, si nous considérons l'action adjointe, l'orbite est la classe de conjugaison : $C_g = G \cdot g$. Et le stabilisateur de g pour l'action adjointe n'est autre que le centralisateur de g :

$$\begin{aligned} \mathrm{Fix}(g) &= \{h \in G \text{ tel que } h \cdot g = g\} && \text{(3.143a)} \\ &= \{h \in G \text{ tel que } hgh^{-1} = g\} && \text{(3.143b)} \\ &= \{h \in G \text{ tel que } gh = hg\} && \text{(3.143c)} \\ &= Z_G(g). && \text{(3.143d)} \end{aligned}$$

Donc la formule $\mathrm{Card}(G \cdot g) = |G : G_g|$ devient, dans le cas de l'action adjointe de G sur lui-même : $\mathrm{Card}(C_g) = |G : Z_G(g)|$. □

Lemme 3.80.
Soit G un groupe agissant sur l'ensemble E. On définit $x \sim x'$ si et seulement s'il existe $g \in G$ tel que $g \cdot x = x'$. Alors

(1) la relation $\sim$ est une relation d'équivalence.

12. Définition 3.2.

(2) la classe $[x]$ est l'orbite $\mathcal{O}_x$ de x sous G.

Corollaire 3.81 (Équation des orbites)**.**
Soit G un groupe agissant sur l'ensemble E et $\mathcal{O}_1, \dots, \mathcal{O}_k$ la liste des orbites (distinctes). Alors

(1) $E = \bigcup_i \mathcal{O}_i$, l'union est disjointe,

(2) $\mathrm{Card}(E) = \sum_i \mathrm{Card}(\mathcal{O}_i)$.

Définition 3.82.
*Soit G un groupe agissant sur l'ensemble E. Un **domaine fondamental** ou une **transversale** est une partie de E contenant un et un seul élément de chaque orbite.*

Autrement dit, les images des éléments d'un domaine fondamental forment une partition de l'ensemble :

$$E = \bigsqcup_{g\in G} g(F) \tag{3.144}$$

où $g(F) = \phi_g(F) = \{\phi_g(x) \text{ tel que } x \in F\}$. L'union est disjointe, c'est à dire que si $g \neq g'$, alors $g(F) \cap g'(F) = \varnothing$.

Proposition 3.83 (Équation des classes[25])**.**
Soit G, un groupe fini opérant sur un ensemble E. Si E'' est un ensemble contenant exactement un élément de chaque orbite dans $E \backslash \mathrm{Fix}_G(E)$, alors

$$|G| = |\mathrm{Fix}_G(E)| + \sum_{x\in E''} \frac{|G|}{|\mathrm{Fix}_G(x)|}. \tag{3.145}$$

Si de plus G est un p-groupe, alors

$$|E| = |\mathrm{Fix}_G(E)| \mod p. \tag{3.146}$$

Démonstration. Par le corollaire 3.81, nous avons $|G| = \sum_{x\in E'} |\mathcal{O}_x|$ où E' est une transversale. En séparant la somme entre les orbites à un élément et les autres,

$$|G| = \mathrm{Card}(\mathrm{Fix}_G(E)) + \sum_{x\in E''} \frac{|G|}{|\mathrm{Fix}_G(x)|} \tag{3.147}$$

où nous avons utilisé le fait que $|G| = |\mathrm{Fix}_G(x)||\mathcal{O}_x|$.

Si G est un p-groupe alors si $x \in E''$, $\mathrm{Fix}_G(x)$ est un sous-groupe propre de G et donc $|\mathrm{Fix}_G(x)|$ est un diviseur propre de $|G|$. Du coup la fraction $|G|/|\mathrm{Fix}_G(x)|$ est une puissance non nulle de p et l'équation (3.145) devient immédiatement (3.146). □

Corollaire 3.84 (Équation des classes)**.**
Soit G, un groupe et $C_1, \dots, C_l$ la liste de ses classes de conjugaison contenant plus de un éléments. Alors

$$\mathrm{Card}(G) = \mathrm{Card}\big(Z(G)\big) + \sum_i |G : Z_{g_i}| = \mathrm{Card}\big(Z(G)\big) + \sum_i \frac{\mathrm{Card}(G)}{\mathrm{Card}\big(\mathrm{Fix}(g_i)\big)} \tag{3.148}$$

si $g_i \in C_i$.

Démonstration. Étant donné que les classes de conjugaison sont disjointes, le cardinal de G est bien la somme des cardinaux de ses classes. Les classes ne contenant que un seul élément sont celles des éléments de $Z(G)$. En ce qui concerne les autres orbites, $\mathrm{Card}(C_{g_i}) = |G : Z_{g_i}|$ par le théorème orbite-stabilisateur (proposition 3.78). □

Théorème 3.85 (Formule de Burnside)**.**
Si G est un groupe fini agissant sur l'ensemble fini E et si Ω est l'ensemble des orbites, alors

$$\mathrm{Card}(\Omega) = \frac{1}{|G|} \sum_{g\in G} \mathrm{Card}\big(\mathrm{Fix}(g)\big). \tag{3.149}$$

Démonstration. Nous considérons l'ensemble

$$A = \{(g, x) \in G \times E \text{ tel que } gx = x\}, \tag{3.150}$$

et nous en calculons le cardinal de deux façons. D'abord

$$\begin{aligned}
\text{Card}(A) &= \sum_{x\in E} \text{Card}\{g \in g \text{ tel que } gx = x\} && (3.151a)\\
&= \sum_{x\in E} \text{Card}(\text{Fix}(x)) && (3.151b)\\
&= \sum_{\omega\in\Omega} \sum_{x\in\omega} \text{Card}(\text{Fix}(x)) && (3.151c)\\
&= \sum_{\omega\in\Omega} \frac{|G|}{\text{Card}(\omega)} && (3.151d)\\
&= |G|. && (3.151e)
\end{aligned}$$

Pour obtenir (3.151d) nous avons utilisé l'équation des classes (3.139). L'autre façon de calculer $\text{Card}(A)$ est de regrouper ainsi :

$$\text{Card}(A) = \sum_{g\in G} \text{Card}\{x \in E \text{ tel que } gx = x\} = \sum_{g\in G} \text{Card}(\text{Fix}(g)). \tag{3.152}$$

En égalisant les deux expressions de $\text{Card}(A)$ nous trouvons

$$|G|\,\text{Card}(\Omega) = \sum_{g\in G} \text{Card}(\text{Fix}(g)). \tag{3.153}$$

□

Proposition 3.86.
Soit G un groupe et H, un sous-groupe du centre de G.

(1) H est normal dans G.

(2) Si G/H est monogène, alors G est abélien.

(3) Si G est fini de centre Z, alors $|G : H|$ n'est pas premier.

Théorème 3.87.
Soit G un groupe cyclique d'ordre n.

(1) Tout sous-groupe de G est cyclique.

(2) Pour chaque diviseur d de n, il existe un unique sous-groupe H_d de G d'ordre d.

Si a est un générateur de G, alors H_d peut être décrit des façons suivantes :

$$H_d = \{x \in G \text{ tel que } x^d = e\} = \{x \in G \text{ tel que } \exists y \in G \text{ tel que } y^{n/d} = x\} = \langle a^{n/d}\rangle. \tag{3.154}$$

Exemple 3.88
Si E est un espace vectoriel alors $(E, +)$ est un groupe commutatif. L'inverse de x est $-x$. △

Définition 3.89.
*Soit G un groupe agissant sur un ensemble E. Nous disons que l'action est **transitive** si elle possède une seule orbite. L'action est **libre** si $g \cdot x = g' \cdot x$ implique $g = g'$.*

3.10 Groupe des permutations

Définition 3.90.
*Soit un ensemble E. Une **permutation** de l'ensemble E est une bijection $E \to E$. Le **groupe symétrique** de E le groupe des bijections $E \mapsto E$; il est noté S_E.*

*Le **groupe symétrique** S_n est le groupe des permutations de l'ensemble $\{1, \dots, n\}$. C'est donc l'ensemble des bijections $\{1, \dots, n\} \to \{1, \dots, n\}$.*

Lemme 3.91 ([26])**.**
Deux résultats.

(1) Tout groupe est isomorphe à une sous-groupe d'un groupe symétrique.

(2) Tout groupe fini d'ordre n est isomorphe à un sous-groupe de S_n.

Démonstration. Soit, pour $g \in G$ donné, l'application

$$\begin{aligned} \tau_g \colon G &\to G \\ x &\mapsto gx. \end{aligned} \tag{3.155}$$

Cela est une bijection de G. En effet $\tau_g(x) = y$ pour $x = g^{-1}y$ (surjectif) et $\tau_g(x) = \tau_g(y)$ impllique $gx = gy$ et donc $x = y$ (injection).

Nous avons donc $\tau_g \in S_G$. De plus l'application

$$\begin{aligned} \varphi \colon G &\to S_G \\ g &\mapsto \tau_g \end{aligned} \tag{3.156}$$

est un morphisme de groupe. Il est injectif parce que si $\tau_g = \tau_h$ alors $gx = hx$ pour tout x. En particulier $g = h$.

Donc $\varphi \colon G \to \text{Image}(\varphi)$ est un isomorphisme entre G et un sous-groupe de S_G.

Un groupe fini de cardinal n est isomorphe à un sous-groupe de S_G ; or S_G est isomorphe à un des S_n. □

3.11 Produit semi-direct de groupes

Définition 3.92.
*Une **suite exacte** est une suite d'applications comme suit :*

$$\cdots \xrightarrow{f_i} A_i \xrightarrow{f_{i+1}} A_{i+1} \xrightarrow{f_{i+2}} \dots \tag{3.157}$$

où pour chaque i, les application f_i et f_{i+1} vérifient $\ker(f_{i+1}) = \text{Image}(f_i)$. Lorsque les ensembles A_i sont des groupes, alors nous demandons de plus que les f_i soient de homomorphismes.

Très souvent nous sommes confrontés à des suites exactes de la forme

$$1 \longrightarrow A \xrightarrow{f} G \xrightarrow{g} B \longrightarrow 1 \tag{3.158}$$

où G, A et B sont des groupes, 1 est l'identité. La première flèche est l'application $\{1\} \to A$ qui à 1 fait correspondre 1. La dernière est l'application $B \to 1$ qui à tous les éléments de B fait correspondre 1. Le noyau de f étant l'image de la première flèche (c'est à dire $\{1\}$), l'application f est injective. L'image de g étant le noyau de la dernière flèche (c'est à dire B en entier), l'application g est surjective.

Définition 3.93.
*Soient N et H deux groupes et un morphisme de groupes $\phi \colon H \to \text{Aut}(N)$. Le **produit semi-direct** de N et H relativement à ϕ, noté $N \times_\phi H$ est l'ensemble $N \times H$ muni de la loi (que l'on vérifiera être de groupe)*

$$(n, h) \cdot (n', h') = (n\phi_h(n'), hh'). \tag{3.159}$$

Attention à l'ordre quelque peu contre intuitif. Lorsque nous notons $N \times_\phi H$, c'est bien $\phi\colon H \to \mathrm{Aut}(N)$, c'est à dire H qui agit sur N et non le contraire.

Lorsque N et H sont des sous-groupes d'un même groupe, le plus souvent ϕ est l'action adjointe définie en 3.76.

Le théorème suivant permet de reconnaître un produit semi-direct lorsqu'on en voit un.

Théorème 3.94 ([7])**.**
Soit une suite exacte de groupes

$$1 \longrightarrow N \xrightarrow{i} G \xrightarrow{s} H \longrightarrow 1 \tag{3.160}$$

S'il existe un sous-groupe $\tilde{H}$ de G à partir duquel s est un isomorphisme, alors

$$G \simeq i(N) \times_\sigma \tilde{H} \tag{3.161}$$

où σ est l'action adjointe[13] *de $\tilde{H}$ sur $i(N)$.*

Démonstration. Nous posons $\tilde{N} = i(N)$ et nous allons subdiviser la preuve en petits pas.

(1) $\tilde{N}$ est normal dans G. En effet étant donné que la suite est exacte nous avons $\tilde{N} = \ker(s)$. Le noyau d'un morphisme est toujours un sous-groupe normal.

(2) $\tilde{N} \cap \tilde{H} = \{e\}$. L'application s étant un isomorphisme depuis $\tilde{H}$, il n'y a pas d'éléments de $\tilde{H}$ dans $\ker(s)$ autre que e.

(3) $G = \tilde{N}\tilde{H}$. Nous considérons $g \in G$ et $h \in \tilde{H}$ tel que $s(g) = s(h)$. L'existence d'un tel h est assurée par le fait que s est surjective depuis $\tilde{H}$. Du coup nous avons $e = s(gh^{-1})$, c'est à dire $gh^{-1} \in \ker(s) = \tilde{N}$. Nous avons donc bien la décomposition $g = (gh^{-1})h$, et donc $G = \tilde{N}\tilde{H}$.

(4) L'écriture $g = nh$ avec $n \in \tilde{N}$ et $h \in \tilde{H}$ est unique. Si $nh = n'h'$, alors $n = n'h'h^{-1}$, ce qui signifierait que $h'h^{-1} \in \tilde{N}$. Mais étant donné que $\tilde{H} \cap \tilde{N} = \{e\}$, nous obtenons $h = h'$ et par suite $n = n'$.

(5) L'application
$$\begin{aligned} \phi\colon G &\to \tilde{N} \times \tilde{H} \\ nh &\mapsto (n, h) \end{aligned} \tag{3.162}$$
est une bijection. C'est une conséquence des points (3) et (4).

(6) Si sur $\tilde{N} \times \tilde{H}$ nous mettons le produit
$$(n, h) \cdot (n', h') = (n\sigma_h n', hh') \tag{3.163}$$
où σ est l'action adjointe du groupe sur lui-même, c'est à dire $\sigma_x(y) = xyx^{-1}$, alors ϕ est un isomorphisme. Si $g, g' \in G$ s'écrivent (de façon unique par le point (5)) $g = nh$ et $g' = n'h'$ alors
$$\begin{aligned} \phi(nhn'h') &= \phi(n \underbrace{hn'h^{-1}}_{\in\tilde{N}} hh') && \text{(3.164a)} \\ &= \phi\big((nhn'h^{-1})(hh')\big) && \text{(3.164b)} \\ &= (nhn'h^{-1}, hh') && \text{(3.164c)} \\ &= (n, h) \cdot (n', h') && \text{(3.164d)} \\ &= \phi(nh)\phi(n'h'). && \text{(3.164e)} \end{aligned}$$

□

Corollaire 3.95.
Soit G un groupe, et N, H des sous-groupes de G tels que

13. Le fait que H agisse sur $i(N)$ fait partie du théorème.

(1) H normalise N (c'est à dire que $hnh^{-1} \in N$ pour tout $h \in H$ et $n \in N$ [14]),

(2) $H \cap N = \{e\}$,

(3) $HN = G$.

Alors l'application

$$\begin{aligned} \psi\colon N \times_\sigma H &\to G \\ (n, h) &\mapsto nh \end{aligned} \tag{3.165}$$

est un isomorphisme de groupes.

Dans les hypothèses, l'ordre entre N et H est important lorsqu'on dit que c'est N qui agit sur H ; mais l'hypothèse $NH = G$ est équivalente à $HN = G$ (passer à l'inverse pour s'en assurer).

Insistons encore un peu sur la notation : dans $N \times_\sigma H$, c'est H qui agit sur N par σ.

3.12 Groupe de torsion

Soit G un groupe. Un élément $g \in G$ est un **élément de torsion** s'il est d'ordre fini. La **torsion** de G est l'ensemble de ses éléments de torsion. Nous disons qu'un groupe est un **groupe de torsion** si tous ses éléments sont de torsion.

Exemple 3.96
Le groupe additif $\mathbb{Q}/\mathbb{Z}$ est un groupe de torsion parce que si $[x] = [p/q]$, alors $q[x] = [p] = [0]$.
△

3.13 Famille presque nulle

Soit $(G, +)$ un groupe abélien et $\mathcal{F} = \{g_i\}_{i\in I}$ une famille d'éléments de G indicés par un ensemble I. Le **support** de $\mathcal{F}$ est l'ensemble $\{i \in I \text{ tel que } g_i \neq 0\}$. La famille est dite **presque nulle** si le support est fini.

Nous disons que $\mathcal{F}$ est une **suite** si $I = \mathbb{N}$.

3.14 Le groupe des racines de l'unité

Définition 3.97.
*Une **racine n^ede l'unité** est une racine du polynôme $X^n - 1$.*

Dans $\mathbb{C}$ nous avons au maximum n telles racines, et il est facile de voir qu'il y en a effectivement n distinctes données par les éléments du groupe multiplicatif

$$\mathrm{U}_n = \{e^{2i\pi k/n} \text{ tel que } k = 0, \dots, n-1\} \tag{3.166}$$

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 5.117.

Lemme 3.98.
L'ensemble U_n est un groupe cyclique [15] 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 $\mathbb{C}$.

14. Ou encore que H agit sur N par automorphismes internes.
15. Définition 3.7.

Le fait que ce soit un groupe cyclique contenant n éléments est la définition. □

Le lemme suivant donne les autres générateurs.

Lemme 3.99.
Le nombre ξ^a est un générateur de U_n si et seulement si $\mathrm{pgcd}(a,n)=1$.

Démonstration. Si $\mathrm{pgcd}(a,n)=1$ alors le théorème de Bézout 3.32 nous fournit des entiers u et v tels que $ua+vn=1$. Alors nous avons

$$e^{2i\pi/n}=e^{2(ua+vn)i\pi/n}=(e^{2ai\pi/n})^u, \tag{3.167}$$

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 $\mathrm{pgcd}(a,n)=1$. □

Exemple 3.100
Une conséquence tout à fait extraordinaire de ce lemme est que 7 est générateur de $\mathbb{Z}/12\mathbb{Z}$ (parce que $\mathrm{pgcd}(7,12)=1$). 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[27]. Ce fait est connu des pianistes [16] depuis des siècles. △

Proposition 3.101 (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

$$(e^{2i\pi k/\alpha})^\beta=e^{2i\pi k\beta/\alpha}=1, \tag{3.168}$$

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

$$k\beta=l\alpha. \tag{3.169}$$

Si $k>0$, comme le nombre α divise $k\beta$, cela conduirait via le lemme de Gauss 3.40 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 3.102 (Intersection : le cas général[1])**.**
Soient des entiers positifs $\alpha_1,\dots,\alpha_p$. Nous avons

$$\bigcup_{i=1}^{p}U_{\alpha_i}=\{1\} \tag{3.170}$$

si et seulement si $\mathrm{pgcd}(\alpha_1,\dots,\alpha_p)=1$ *(c'est à dire que les α_i sont premiers dans leur ensemble).*

16. Même ceux qui ignorent le théorème de Bézout.

Démonstration. Nous le décomposons les α_i en facteurs premiers[17] de la façon suivante : $\alpha_i = \prod_{k\in\mathbb{N}} p_k^{\alpha_i^{(k)}}$ où les p_k sont les nombres premiers.

Caractérisation par une décomposition en facteurs premiers Les éléments z différents de 1 dans U_{α_1} s'écrivent sous la forme

$$z = e^{2i\pi k/\alpha_1} \tag{3.171}$$

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

$$\alpha_1^{(l)} - \alpha_i^{(l)} \tag{3.172}$$

pour tout l.

Sens direct L'hypothèse $\operatorname{pgcd}(\alpha_1, \ldots, \alpha_p) \neq 1$ 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 $\operatorname{pgcd}(\alpha_1, \ldots, \alpha_p) = 1$ 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 3.103.
Les générateurs de U_n *sont les* ***racines primitives***[18] *de l'unité dans* $\mathbb{C}$*. Nous nommons* Δ_n *leur ensemble :*

$$\Delta_n = \{e^{2ki\pi/n} \text{ tel que } 0 \leqslant k \leqslant n-1, \operatorname{pgcd}(k,n) = 1\}. \tag{3.173}$$

Nous avons par exemple

$$\Delta_1 = \{1\} \tag{3.174a}$$
$$\Delta_2 = \{e^{\pi i}\} \tag{3.174b}$$
$$\Delta_4 = \{e^{\pi i/2}, e^{3\pi i/2}\}. \tag{3.174c}$$

Notons que $1 \in \Delta_d$ seulement avec $d = 1$.

3.15 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 8.2.5.

17. Théorème 3.42.

18. 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 5.154.

3.15.1 Introduction par les racines de l'unité

Définition 3.104.
La fonction φ donnée par

$$\varphi(n) = \mathrm{Card}(\Delta_n) \tag{3.175}$$

*est l'**indicatrice d'Euler**.*

Si p est un nombre premier, alors $\varphi(p) = p - 1$.

Lemme 3.105.
Nous avons

$$\mathrm{U}_n = \bigcup_{d|n} \Delta_d \tag{3.176}$$

et l'union est disjointe. Nous avons aussi la formule

$$n = \sum_{d|n} \varphi(d). \tag{3.177}$$

Démonstration. À l'application $x \mapsto e^{2i\pi x}$ près, nous pouvons considérer

$$\Delta_d = \{\frac{k}{d} \text{ tel que } k = 0, \dots, d-1, \mathrm{pgcd}(k,d) = 1\}, \tag{3.178}$$

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

$$\mathrm{U}_n = \{\frac{k}{n} \text{ tel que } k = 0, \dots, n-1\}. \tag{3.179}$$

L'égalité (3.176) 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 (3.177) consiste à prendre le cardinal des deux côtés de (3.176). 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

$$\Phi_n(d) = \{m \in \mathbb{N} \text{ tel que } \mathrm{pgcd}(m,n) = \frac{n}{d}\}. \tag{3.180}$$

Étant donné que tous les entiers entre 1 et n ont un pgcd avec n qui est automatiquement un quotient de n nous avons

$$\{1, \dots, n\} = \bigcup_{d|n} \Phi_n(d) \tag{3.181}$$

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

$$\mathrm{Card}\{1, \dots, n\} = \sum_{d|n} \mathrm{Card}(\Phi_n(d)) = \sum_{d|n} \varphi(d). \tag{3.182}$$

□

Lemme 3.106.
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 $q \leqslant p^{n-1}$ [19]. Il y a évidemment p^{n-1} tels nombres.

Par conséquent le cardinal de P_{p^n} est $\varphi(p^n) = p^n - p^{n-1}$. □

19. Corollaire 3.44.

Problèmes et choses à faire

P_n n'a pas été défini.

Définition proposée (et vue par après) : $P_n = \{m \in \mathbb{N} \text{ tel que } \operatorname{pgcd}(m, n) = 1\}$. À mettre donc en lien avec Δ_d.

3.15.2 Générateurs

Proposition 3.107.
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 $\operatorname{gr}\big([1]_n\big) = \mathbb{Z}/n\mathbb{Z}$. 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

$$ux + vn = 1. \tag{3.183}$$

Cela signifie entre autres que[20] $x\mathbb{Z} + n\mathbb{Z} = \mathbb{Z}$, et aussi que $\operatorname{pgcd}(x, n) = 1$ par le théorème de Bézout 3.32, et donc que $x \in P_n$. □

20. Corollaire 3.33

Chapitre 4

Anneaux

4.1 Généralités

On rappelle la définition 2.20 d'un anneau ainsi que les notations 0 pour le neutre additif et 1 pour le neutre multiplicatif. Certains ouvrages ne demandent pas spécialement de neutre pour la multiplication, et ajoutent le terme unitaire pour spécifier qu'il y a un neutre.

Lemme 4.1.
Pour tout élément a d'un anneau nous avons $a \cdot 0 = 0$.

Démonstration. L'élément 0 est le neutre de l'addition. Il peut être écrit $1 - 1$, et en utilisant la distributivité,

$$a \cdot 0 = a \cdot (1 - 1) = a - a = 0. \tag{4.1}$$

Notons que la dernière égalité s'écrit en détail $a-a = a+(-a)$ qui donne le neutre de l'addition. □

Proposition 4.2.
Dans un anneau non-nul, le neutre pour l'addition est distinct du neutre pour la multiplication.

Démonstration. Supposons par contraposée que dans un anneau A, $1 = 0$. Alors, pour tout $a \in A$, on a $a = 1a = 0a = (1 - 1)a = a - a = 0$, d'où l'on déduit $-a = 0$ et par suite, $a = 0$. □

Soit X un ensemble et un anneau $(A, +, \times)$. Nous considérons $\mathrm{Fun}(X, A)$ l'ensemble des applications $X \to A$. Cet ensemble devient un anneau avec les définitions

$$(f + g)(x) = f(x) + g(x) \tag{4.2a}$$
$$(fg)(x) = f(x)g(x). \tag{4.2b}$$

Cela est la **structure canonique** d'anneau sur $\mathrm{Fun}(X, A)$.

Le **centralisateur** de $x \in A$ dans A est l'ensemble

$$\{y \in A \text{ tel que } xy = yx\}, \tag{4.3}$$

le **centre** de A est

$$\{y \in A \text{ tel que } xy = yx, \forall x \in A\}. \tag{4.4}$$

Définition 4.3 (Diviseurs dans un anneau)**.**
*Soient $a, b \in A$. On dit que a divise b, ou que a est un **diviseur (à gauche)** de b s'il existe $c \in A$ tel que $ac = b$. On dit que c'est un diviseur de b à droite si $ca = b$ pour un certain $cinA$.*

Un cas particulier est le cas des diviseurs de zéro. L'absence de tels diviseurs dans un anneau est une propriété intéressante : on dit dans ce cas que l'anneau est intègre. Nous étudions ces anneaux plus en détail en section 4.6.

Un élément $a \in A$ est **régulier à droite** si $ba = 0$ implique $b = 0$. Il est régulier à gauche si $ab = 0$ implique $b = 0$.

Définition 4.4 (Éléments nilpotents, unipotents et inversibles)**.**
*On dit que $a \in A$ est **nilpotent** s'il existe $n \in \mathbb{N}$ tel que $a^n = 0$. Il est dit **unipotent** si $a - 1$ est nilpotent, c'est à dire si $(a-1)^n = 0$ pour un certain $n \in \mathbb{N}$.*

*Un élément $a \in A$ est dit **inversible** s'il existe $b \in A$ tel que $ab = 1$.*

L'ensemble $U(A)$ des éléments inversibles de A est un groupe pour la multiplication. Nous notons $A^* = A\backslash\{0\}$.

La proposition suivante donne une caractérisation d'un corps, en disant un tout petit peu plus que la définition 2.21.

Proposition 4.5.
L'anneau A est un corps si et seulement si $U(A) = A^*$.

Démonstration. L'inclusion $A^* \subset U(A)$ est la définition.

Pour l'inclusion inverse, il faut montrer que les inversibles ne peuvent pas être zéro, c'est à dire que zéro n'est pas inversible. Cela n'est autre que le lemme 4.1 couplé à la proposition 4.2 : $a \cdot 0 = 0 \neq 1$ pour tout a. □

Lemme 4.6.
Si a et b commutent, nous avons, pour tout $r \in \mathbb{N}$ et $r > 0$, la formule

$$a^{r+1} - b^{r+1} = (a-b)\left(\sum_{k=0}^{r} a^{r-k}b^k\right). \tag{4.5}$$

Démonstration. Démontrons cela par récurrence. Le cas $r = 0$ est évident. Pour un r donné, si (4.5) est vraie, alors

$$\begin{aligned}
a^{r+2} - b^{r+2} &= a^{r+1}a - a^{r+1}b + a^{r+1}b - b^{r+1}b \\
&= a^{r+1}(a-b) + (a^{r+1} - b^{r+1})b \\
&= a^{r+1}(a-b) + (a-b)\left(\sum_{k=0}^{r} a^{r-k}b^k\right)b \\
&= (a-b)\left(a^{r+1} + \left(\sum_{k=0}^{r} a^{r-k}b^k\right)b\right) \\
&= (a-b)\left(a^{r+1} + \sum_{k=0}^{r} a^{r-k}b^{k+1}\right) \\
&= (a-b)\left(a^{r+1} + \sum_{k'=1}^{r+1} a^{(r+1)-k'}b^{k'}\right) \\
&= (a-b)\left(\sum_{k'=0}^{r+1} a^{(r+1)-k'}b^{k'}\right).
\end{aligned}$$

□

Proposition 4.7.
Si a est un élément nilpotent de l'anneau A, alors $1 - a$ est inversible. Si a est nilpotent non nul, alors il est diviseur de zéro.

Démonstration. Soit n le minimum tel que $a^n = 0$. En vertu de la formule (4.5) nous avons

$$1 = 1 - a^n = (1-a)(1 + a + \cdots + a^{n-1}) = (1 + a + \cdots + a^{n-1})(1-a). \tag{4.6}$$

La somme $1 + a + \cdots + a^{n-1}$ est donc un inverse de $(1-a)$. □

Définition 4.8.
Soit A un anneau et $a, b \in A$. Nous disons que d est un pgcd *de a et b si tout diviseur commun de a et b divise d.*

Nous rappelons également la définition 2.23 de morphisme d'anneaux. Remarquons que si f est un morphisme, nous avons $f(0) = 0$ et $f(x)^{-1} = f(x^{-1})$.

4.2 Binôme de Newton et morphisme de Frobenius

Proposition 4.9.
Pour tout x, $y \in \mathbb{R}$ et $n \in \mathbb{N}$, nous avons

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \tag{4.7}$$

où

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{4.8}$$

sont les ***coefficients binomiaux****.*

La preuve qui suit provient de wikipédia.

Démonstration. La preuve se fait par récurrence. La vérification pour $n = 0$ et $n = 1$ se fait aisément pour peu que l'on se rappelle que $x^0 = 1$ et que $0! = 1$, ce qui donne entre autres $\binom{0}{0} = 1$.

Supposons que la formule (4.7) soit vraie pour $n \geqslant 1$, et prouvons la pour $n+1$. Nous avons

$$\begin{aligned}(x+y)^{n+1} &= (x+y) \cdot \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \\ &= \sum_{k=0}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^{k+1} \\ &= x^{n+1} + \sum_{k=1}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^{n-1} \binom{n}{k} x^{n-k} y^{k+1} + y^{n+1}.\end{aligned} \tag{4.9}$$

La seconde grande somme peut être transformée en posant $i = k+1$:

$$\sum_{k=0}^{n-1} \binom{n}{k} x^{n-k} y^{k+1} = \sum_{i=1}^{n} \binom{n}{i-1} x^{n-(i-1)} y^{i-1+1}, \tag{4.10}$$

dans lequel nous pouvons immédiatement renommer i par k. En remplaçant dans la dernière expression de (4.9), nous trouvons

$$(x+y)^{n+1} = x^{n+1} + y^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k} + \binom{n}{k-1} \right] x^{n-k+1} y^k. \tag{4.11}$$

La thèse découle maintenant de la formule

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k} \tag{4.12}$$

qui est vraie parce que

$$\frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)(n-k+1)!} = \frac{n!(n-k+1) + n!k}{k!(n-k+1)!} = \frac{n!(n+1)}{k!(n-k+1)!}, \tag{4.13}$$

par simple mise au même dénominateur. □

4.3 Idéal dans un anneau

La définition d'un idéal dans un anneau est la définition 2.24.

Définition 4.10 (Idéal engendré par un élément)**.**
*Si p est un élément d'un anneau A alors nous notons (p) l'idéal dans A **engendré** par p, c'est à dire pA.*

Définition 4.11.
*Un sous ensemble $B \subset A$ d'un anneau est un **sous anneau** si*

(1) $1 \in B$

(2) B *est un sous-groupe pour l'addition*

(3) B *est stable pour la multiplication.*

Remarque 4.12.
Un idéal n'est pas toujours un anneau parce que l'identité pourrait manquer. Un idéal qui contient l'identité est l'anneau complet.

Exemple 4.13
L'ensemble $2\mathbb{Z}$ est un idéal de $\mathbb{Z}$. On peut aussi le noter (2). △

Soit A, un anneau, I un idéal bilatère[1] de A. Nous considérons la relation d'équivalence $x \sim y$ si et seulement si $x - y \in I$. Dans ce cas, le quotient

$$A/\sim = A/I \tag{4.14}$$

est un anneau appelé **anneau quotient**. La surjection $A \to A/I$ est un morphisme.

Proposition 4.14 (Premier théorème d'isomorphisme pour les anneaux)**.**
Soient A et B des anneaux et un homomorphisme $f\colon A \to B$. Nous considérons l'injection canonique $j\colon f(A) \to B$ et la surjection canonique $\phi\colon A \to A/\ker f$. Alors il existe un unique isomorphisme

$$\tilde{f}\colon A/\ker f \to f(A) \tag{4.15}$$

tel que $f = j \circ \tilde{f} \circ \phi$.

$$\begin{array}{ccc} A & \xrightarrow{\quad f \quad} & B \\ {\scriptstyle\phi}\downarrow & & \downarrow{\scriptstyle j} \\ A/\ker f & \xrightarrow[\tilde{f}]{} & f(A) \subset B \end{array} \tag{4.16}$$

Proposition 4.15.
Soit I, un idéal de A et $\phi\colon A \to A/I$ la surjection canonique. Les idéaux de A/I sont les $\phi(J)$ où J est un idéal de A contenant I. De plus cette relation est bijective :

$$\{\textit{idéaux de } A \textit{ contenant } I\} \simeq \{\textit{idéaux de } A/I\}. \tag{4.17}$$

Démonstration. Si $I \subset J$ et si J est un idéal de A, alors $\phi(J)$ est un idéal dans A/I. En effet un élément de $\phi(J)$ est de la forme $\phi(j)$ et un élément de A/I est de la forme $\phi(i)$. Leur produit vaut

$$\phi(i)\phi(j) = \phi(ij) \in \phi(J). \tag{4.18}$$

Soit maintenant K, un idéal dans A/I. Soit $J = \phi^{-1}(K)$. Étant donné qu'un idéal doit contenir 0 (parce qu'un idéal est un groupe pour l'addition), $[0] \in K$ et par conséquent $I \subset \phi^{-1}(K)$. □

1. Définition 2.24.

4.3.1 Résultats supplémentaires sur l'anneau des entiers

Corollaire 4.16.
Les quotients de $\mathbb{Z}$ sont $\mathbb{Z}/n\mathbb{Z}$.

Démonstration. Tous les idéaux de $\mathbb{Z}$ sont de la forme $n\mathbb{Z}$. En effet en vertu de la proposition 3.27, les seuls sous-groupes de $\mathbb{Z}$ (en tant que groupe additif) sont les $n\mathbb{Z}$. Tous les idéaux sont donc de cette forme. De plus les $n\mathbb{Z}$ sont effectivement tous des idéaux : si $a \in n\mathbb{Z}$ et si $k \in \mathbb{Z}$ alors $ak \in n\mathbb{Z}$. Cela est donc un idéal. □

Proposition 4.17.
Soit $n \geqslant 2$ un entier et $\phi\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, la surjection canonique. Nous noterons $\tilde{a} = \phi(a)$. Alors l'ensemble des inversibles de $\mathbb{Z}/n\mathbb{Z}$ est donné par

$$U(\mathbb{Z}/n\mathbb{Z}) = \phi(P_n) = \{\tilde{x} \text{ tel que } 0 \leqslant x \leqslant n \text{ tel que } \operatorname{pgcd}(x, n) = 1\}. \tag{4.19}$$

où P_n est l'ensemble $P_n = \{x \in \{0, \dots, n-1\} \text{ tel que } \operatorname{pgcd}(x, n) = 1\}$. En particulier, $\operatorname{Card}\big(U(\mathbb{Z}/n\mathbb{Z})\big) = \varphi(n)$.

Démonstration. Soit $0 \leqslant x \leqslant n$ tel que $\operatorname{pgcd}(x, n) = 1$. Il existe donc [2] $u, v \in \mathbb{Z}$ tels que $ux+vn = 1$. En passant aux classes,

$$\tilde{u}\tilde{x} = \tilde{1}, \tag{4.20}$$

donc $\tilde{u}$ est l'inverse de $\tilde{x}$. Cela prouve que $\phi(P_n) \subset U(\mathbb{Z}/n\mathbb{Z})$.

Nous prouvons maintenant l'inclusion inverse. Soit $\tilde{x}$ et $\tilde{y}$ inverses l'un de l'autre : $\tilde{x}\tilde{y} = \tilde{1}$. Il existe donc $q \in \mathbb{Z}$ tel que $xy - qn = 1$, ce qui prouve [3] que $\operatorname{pgcd}(x, n) = 1$. □

4.4 Caractéristique

Lemme-définition 4.18.
Soit l'application

$$\begin{aligned} \mu\colon \mathbb{Z} &\to A \\ n &\mapsto n \cdot 1_A. \end{aligned} \tag{4.21}$$

(1) C'est un morphisme d'anneaux.

(2) Le noyau est un sous-groupe de $\mathbb{Z}$

(3) Il existe un unique $p \in \mathbb{Z}$ tel que $\ker(\mu) = p\mathbb{Z}$.

Ce p est la ***caractéristique*** *de A.*

Par exemple la caractéristique que $\mathbb{Q}$ est zéro parce qu'aucun multiple de l'unité n'est nul.
À propos de diagonalisation en caractéristique 2, voir l'exemple 9.147.

Lemme 4.19.
Si A est de caractéristique nulle, alors A est infini.

Démonstration. En effet, $\ker \mu = \{0\}$ implique que $n1_A \neq m1_A$ dès que $n \neq m$ et par conséquent A contient $\mathbb{Z}1_A$, et est infini. □

Lemme 4.20.
Si p est la caractéristique de l'anneau A, alors nous avons l'isomorphisme d'anneaux

$$\mathbb{Z}1_A \simeq \mathbb{Z}/p\mathbb{Z}. \tag{4.22}$$

Démonstration. L'isomorphisme est donné par l'application $n1_A \mapsto \phi(n)$ si ϕ est la projection canonique $\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$. □

2. Théorème de Bézout 3.32
3. À nouveau avec le Théorème de Bézout.

Proposition 4.21.
La caractéristique d'un anneau fini divise son cardinal.

Démonstration. Si A est un anneau, le groupe $\mathbb{Z}$ agit sur A par

$$n \cdot a = a + n1_A. \tag{4.23}$$

Chaque orbite de cette action est de la forme

$$\mathcal{O}_a = \{a + n1_A \text{ tel que } n = 0, \ldots, p-1\} \tag{4.24}$$

où p est la caractéristique de A. Les orbites ont p éléments et forment une partition de A, donc le cardinal de A est un multiple de p. □

Lemme 4.22 ([28])**.**
Un anneau totalement ordonné est de caractéristique nulle.

Démonstration. Le morphisme $\mu\colon \mathbb{Z} \to A$, $n \mapsto n1_A$ est strictement croissant, en particulier $\mu(x) \neq \mu(y)$ dès que $x \neq y$. Donc $\ker(\mu) = \{0\}$. □

L'ensemble typique de caractéristique p est $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Proposition 4.23.
Soit A un anneau commutatif de caractéristique première p. Alors $\sigma(x) = x^p$ est un automorphisme de l'anneau A. Nous avons la formule

$$(a+b)^p = a^p + b^p \tag{4.25}$$

pour tout $a, b \in A$.

Démonstration. Nous utilisons la formule du binôme de la proposition 4.9 et le fait que les coefficients binomiaux non extrêmes sont divisibles par p et donc nuls. □

Proposition 4.24.
Soit A un anneau commutatif unitaire de caractéristique p. L'application

$$\begin{aligned} \mathrm{Frob}_A\colon A &\to A \\ x &\mapsto x^p \end{aligned} \tag{4.26}$$

est un automorphisme d'anneau unitaire.

Nous le nommons le **morphisme de Frobenius**. Nous utiliserons aussi les itérés du morphisme de Frobenius : $\mathrm{Frob}^k\colon x \mapsto x^{p^k}$.

Exemple 4.25
Soit à factoriser $X^p - 1$ dans $\mathbb{F}_p$. Grâce au morphisme de Frobenius, nous avons immédiatement

$$X^p - 1 = (X-1)^p. \tag{4.27}$$

△

4.5 Modules

Si A est un anneau et si $(M, +)$ est un groupe commutatif, nous disons que M est un **module** à gauche sur A si nous avons une application $A \times M \to M$ notée $a \cdot x$ telle que

(1) $\alpha \cdot (x+y) = a \cdot x + a \cdot y$ (distributivité de $\cdot$ par rapport à l'addition dans M)

(2) $(a+b)\cdot x = a\cdot x + b\cdot x$ (distributivité de $\cdot$ par rapport à l'addition dans $\mathbb{A}$). Remarque : la loi + du membre de gauche est celle de l'anneau A et la loi + du membre de droite est celle du groupe M

(3) $(ab)\cdot x = a\cdot(b\cdot x)$

(4) $1\cdot x = x$

Remarque 4.26.
La notion de module généralise celle d'espace vectoriel que nous verrons plus loin.

Soit M un A-module et $x = (x_i)_{i\in I}$ une famille d'éléments de M, paramétrée par l'ensemble I. Nous considérons l'application

$$\begin{aligned}\mu_x\colon A^{(I)} &\to M\\ (a_i)_{i\in I} &\mapsto \sum_{i\in I} a_i x_i.\end{aligned} \tag{4.28}$$

Ici $A^{(I)}$ désigne l'ensemble de toutes les applications $I \to A$ de support fini.

Définition 4.27.
À l'instar des espaces vectoriels, les modules ont une notion de partie libre, génératrice et de bases :

*(1) Si μ_x est surjective, nous disons que x est une partie **génératrice**.*

*(2) Si μ_x est injective, nous disons que la partie x est **libre**.*

*(3) Si μ_x est bijective, nous disons que la partie x est une **base**.*

Définition 4.28.
*Un sous-ensemble $N \subset M$ est un **sous-module** si $(N,+)$ est un sous-groupe de $(M,+)$ et si $a\cdot x \in N$ pour tout $x \in N$ et pour tout $a \in A$.*

Exemple 4.29
Un anneau A est lui-même un A-module et ses sous-modules sont les idéaux. △

Définition 4.30.
*Soit M un module sur un anneau commutatif A. Un **projecteur** est une application linéaire $p\colon M \to M$ telle que $p^2 = p$.*

*Une famille $(p_i)_{i\in I}$ sur M est **orthogonale** si $p_i \circ p_j = 0$ pour tout $i \neq j$. La famille est **complète** si $\sum_{i\in I} p_i = \mathbb{1}$.*

Théorème 4.31.
Soient des sous modules $M_1,\dots,M_n$ du module M tels que $M = M_1 \oplus \dots \oplus M_n$. Les applications p_i définies par

$$p_i(x_1 + \dots + x_n) = x_i \tag{4.29}$$

forment une famille orthogonale de projecteurs et $p_1 + \dots + p_n = \mathrm{Id}$.

Inversement, si $(p_1,\dots,p_n)$ est une famille orthogonale de projecteurs dans un module $\mathcal{E}$ tel que $\sum_{i=1}^n p_i = \mathrm{Id}$, alors

$$M = \bigoplus_{i=1}^{n} p_i(M). \tag{4.30}$$

Un module est **simple** ou **irréductible** s'il n'a pas d'autre sous-modules que $\{0\}$ et lui-même. Un module est **indécomposable** s'il ne peut pas être écrit comme somme directe de sous-modules.

Un module simple est a fortiori indécomposable. L'inverse n'est pas vrai comme le montre l'exemple suivant.

Exemple 4.32
Soit $\mathcal{E} = \mathbb{C}[X]/(X^2)$ vu comme $\mathbb{C}[X]$-module. C'est le $\mathbb{C}[X]$-module des polynômes de la forme $aX + b$ avec $a, b \in \mathbb{C}$. L'ensemble des polynômes de la forme aX est un sous-module. Le module

$\mathcal{E}$ n'est donc pas simple. Il est cependant indécomposable parce que $\{aX\}$ est le seul sous-module non trivial. En effet si $\mathcal{F}$ est un sous-module de $\mathcal{E}$ contenant $aX + b$ avec $b \neq 0$, alors $\mathcal{F}$ contient $X(aX + b) = bX$ et donc contient tout $\mathcal{E}$. △

Définition 4.33 (Algèbre[29])**.**
Si $\mathbb{K}$ est un corps commutatif[4]*, une $\mathbb{K}$-**algèbre** A est un espace vectoriel muni d'une opération bilinéaire $\times : A \times A \to A$, c'est à dire telle que pour tout $x, y, z \in A$ et pour tout $\alpha, \beta \in \mathbb{K}$,*

(1) $(x + y) \times z = x \times z + y \times z$

(2) $x \times (y + z) = x \times y + x \times z$

(3) $(\alpha x) \times (\beta y) = (\alpha\beta)(x \times y)$.

*Si A et B sont deux $\mathbb{K}$-algèbres, une application $f : A \to B$ est un **morphisme d'algèbres** entre A et B si pour tout $x, y \in A$ et pour tout $\alpha \in \mathbb{K}$,*

(1) $f(xy) = f(x)f(y)$

(2) $f(x + \alpha y) = f(x) + \alpha f(y)$

où nous avons noté xy pour $x \times y$.

Lemme 4.34 ([1])**.**
Soit une algèbre A et une famille $(X_i)_{i\in I}$ de sous-algèbres de A (ici I est un ensemble quelconque). Alors la partie $X = \bigcap_{i\in I} X_i$ est une sous-algèbre de A.

Démonstration. Nous devons prouver que si x et y sont dans X et $\lambda \in \mathbb{K}$, alors xy, $x + y$ et λx sont dans X. Pour tout $i \in I$ nous avons $x, y \in X_i$ et donc $xy \in X_i$, $x + y \in X_i$ et $\lambda x \in X_i$ (parce que X_i est une algèbre). Donc xy,$x + y$ et λx sont dans X_i pour tout I, et donc dans X. □

Définition 4.35.
*L'**algèbre engendrée** par X est l'intersection de toutes les sous-algèbres de A contenant X (qui est une algèbre par le lemme 4.34).*

4.6 Anneau intègre

Rappelons la notion de diviseurs que nous avons fixée par la définition 4.3. Un élément $a \neq 0$ est un **diviseur de zéro à gauche** s'il existe $x \neq 0$ tel que $xa = 0$. L'élément a est un diviseur de zéro **à droite** s'il existe b tel que $ab = 0$.

Définition 4.36.
*Un anneau est **intègre** s'il est non nul et ne possède pas de diviseurs de zéro.*

Proposition 4.37.
Soit A un anneau. Les assertions suivantes sont équivalentes :

(1) A est un anneau intègre.

(2) La règle du produit nul s'applique dans A : pour tous $a, b \in A$, si $ab = 0$, alors $a = 0$ ou $b = 0$.

(3) On peut simplifier par un même élément non-nul, deux expressions produit dans A qui sont égales : pour tous $a, b, c \in A$ avec $a \neq 0$, si $ab = ac$, alors $b = c$.

Un corps dans lequel l'élément nul est inversible est appelé un mensonge, comme le prouve le lemme suivant.

Lemme 4.38.
En tant que anneau, un corps est un anneau intègre : il n'a pas de diviseurs de zéro.

4. Définition 2.21

Démonstration. En effet si a est un diviseur de zéro, alors $ax = 0$ pour un certain $x \neq 0$. Si a était inversible, nous aurions $x = a^{-1}ax = 0$, ce qui est impossible. □

Conséquence : dans un corps nous avons toujours la règle du produit nul.

Exemple 4.39
L'ensemble $\mathbb{Z}$ avec les opérations usuelles est un anneau intègre. △

Exemple 4.40
L'anneau $\mathbb{Z}/6\mathbb{Z}$ n'est pas intègre parce que $3 \cdot 2 = 0$ alors que ni 3 ni 2 ne sont nuls. △

Exemple 4.41
Nous verrons au théorème 4.103 que si A est intègre, alors l'anneau des polynômes sur A est intègre, et inversement. Nous pouvons néanmoins déjà montrer que le produit de deux polynômes P et Q non-nuls à coefficients dans un anneau intègre [5] est aussi non-nul. En effet, le coefficient du terme de plus haut degré de PQ est donné par le produit des coefficients des termes plus haut degré de P et Q, produit qui est non nul parce que A est intègre. △

4.6.1 Caractéristique d'un anneau intègre

Corollaire 4.42.
L'anneau $\mathbb{Z}/n\mathbb{Z}$ est intègre si et seulement si n est premier.

Démonstration. Si n est premier, tous les éléments de $\mathbb{Z}/n\mathbb{Z}$ sont inversibles parce que tous les éléments rentrent dans $\phi(P_n)$. Donc $\mathbb{Z}/n\mathbb{Z}$ est intègre.

Si n n'est pas premier, il existe $p, q \in \mathbb{N}^*$ tels que $pq = n$. Dans ce cas au niveau des classes nous avons $\tilde{p}\tilde{q} = 0$ avec $\tilde{p} \neq 0 \neq \tilde{q}$, ce qui montre que $\mathbb{Z}/n\mathbb{Z}$ a des diviseurs de zéro et n'est pas intègre. □

Lemme 4.43.
La caractéristique [6] d'un anneau intègre est zéro ou un nombre premier.

Démonstration. Si A est intègre, alors $\mathbb{Z}1_A$ est intègre (a fortiori), et soit A est de caractéristique nulle (auquel cas c'est fini), soit A est de caractéristique p, et $\mathbb{Z}/p\mathbb{Z}$ est intègre parce qu'il est isomorphe à $\mathbb{Z}1_A$. Mais nous savons que $\mathbb{Z}/p\mathbb{Z}$ est intègre si et seulement si p est premier (proposition 4.42). □

Exemple 4.44
Il existe des corps dont la caractéristique n'est pas égale au cardinal (contrairement à ce que laisserait penser l'exemple des $\mathbb{Z}/p\mathbb{Z}$). En effet les matrices $n \times n$ inversibles sur $\mathbb{F}_3$ forment un corps qui n'est pas de cardinal trois alors que la caractéristique est 3 :

$$\begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} + \begin{pmatrix} 1 & \\ & 1 \end{pmatrix} = 0. \tag{4.31}$$

△

Exemple 4.45
Si $\mathbb{K}$ est un corps de caractéristique 2, alors l'égalité $x = -x$ n'implique pas $x = 0$, vu que $2x = 0$ est vérifiée pour tout x. Cela se répercute sur un certain nombre de résultats. Par exemple, en caractéristique deux, une forme antisymétrique n'est pas toujours alternée : voir le lemme 9.2. △

5. Si quelqu'un connais une application de l'utilisation de polynômes à coefficients dans un anneau pas intègre, qu'il me le dise.

6. Définition 4.18.

4.6.2 Divisibilité et classes d'association

Lemme 4.46.
Si A est un anneau intègre et si $a, b \in A$ sont tels que $a \mid b$ et $b \mid a$, alors il existe un inversible $u \in A$ tel que $a = ub$.

Démonstration. Les hypothèses à propos de la divisibilité nous indiquent que $a = xb$ et $b = ya$ pour certains $x, y \in A$. Du coup,

$$b(1 - yx) = 0. \tag{4.32}$$

Étant donné que $\mathbb{A}$ est intègre, cela montre que $b = 0$ ou $1 - yx = 0$. Si $b = 0$ nous avons immédiatement $a = 0$ et le lemme est prouvé. Si au contraire $yx = 1$, c'est que y et x sont inversibles et inverses l'un de l'autre. □

Définition 4.47.
*On dit de deux éléments $a, b \in A$ qu'ils sont **associés** s'ils vérifient les hypothèses du lemme 4.46 ; en d'autres termes, a et b sont associés s'il existe un inversible $u \in A$ tel que $a = ub$.*

*La **classe d'association** d'un élément $a \in A$ est l'ensemble des éléments qui lui sont associés ; en d'autres termes, c'est $aU(A)$.*

Exemple 4.48
Dans $\mathbb{Z}[i]$, les inversibles sont ± 1 et $\pm i$. Donc les éléments associés à z sont z, $-z$, iz et $-iz$.

Notons au passage que la notion de divisibilité dans $\mathbb{Z}[i]$ n'est pas immédiatement intuitive. En effet bien que 7 ne soit pas divisible par 2 (ni dans $\mathbb{Z}$ ni dans $\mathbb{Z}[i]$), le nombre $7 + 6i$ est divisible par $2 + i$ dans $\mathbb{Z}[i]$. En effet :

$$0(2 + i)(4 + i) = 7 + 6i. \tag{4.33}$$

△

Problèmes et choses à faire

Est-ce que quelqu'un connais un anneau contenant $\mathbb{Z}$ dans lequel 7 est divisible en 2 ?

Peut-être $\mathbb{Z}$ étendu par tous les $1/2^n$?

4.6.3 PGCD et PPCM

Pour le théorème de Bézout et autres opérations avec des molulo, voir le thème 47.

Commençons par donner une autre vision de la divisibilité dans les anneaux intègres.

Proposition 4.49.
Dans un anneau intègre A, on a l'équivalence suivante concernant deux éléments $a, b \in A$:

$$a \mid b \Leftrightarrow (b) \subset (a). \tag{4.34}$$

Donc la divisibilité devient en réalité une relation d'ordre dont nous pouvons chercher un maximum et un minimum. Si S est une partie de A, nous notons $a \mid S$ pour exprimer que $a \mid x$ pour tout $x \in S$; de la même façon, $S \mid b$ signifie que $x \mid b$ pour tout $x \in S$

Définition 4.50.
*Soit A, un anneau intègre et $S \subset A$. Nous disons que $\delta \in A$ est un **PGCD** de S si*

(1) $\delta \mid S$

(2) si $d \mid S$ alors[7] $d \mid \delta$.

*Nous disons que $\mu \in A$ est un **PPCM** de S si*

(1) $S \mid \mu$,

(2) si $S \mid m$, alors $\mu \mid m$.

Notons qu'il n'y a en général pas unicité du PGCD ou du PPCM d'un ensemble.

7. Il me semble qu'à ce niveau il y a une faute de frappe dans [30].

Lemme 4.51.
Soit A un anneau intègre et $S \subset A$. Si δ est un PGCD de S, alors l'ensemble des PGCD de S est la classe d'association de δ.

De la même façon si μ est un PPCM de S, alors l'ensemble des PPCM de S est la classe d'association de μ.

Démonstration. Soit δ un PGCD de S et u un inversible dans A. Si $x \in S$ nous avons $\delta \mid x$ et donc $x = a\delta$. Par conséquent $x = au^{-1}u\delta$ et donc $u\delta$ divise x. De la même manière, si d divise x pour tout $x \in S$, alors d divise δ et donc $\delta = ad$ et $u\delta = aud$, ce qui signifie que d divise $u\delta$.

Dans l'autre sens nous devons prouver que si δ' est un autre PGCD de S, alors il existe un inversible $u \in \mathbb{A}$ tel que $\delta' = u\delta$. Vu que δ' divise x pour tout $x \in S$, nous avons $\delta' \mid \delta$, et symétriquement nous trouvons $\delta \mid \delta'$. Par conséquent (lemme 4.46), il existe un inversible u tel que $\delta = u\delta'$.

Le même type de raisonnement tient pour le PPCM. □

Si δ est un PGCD de S, nous dirons *par abus de langage* que δ est *le* PGCD de S, gardant en tête qu'en réalité toute sa classe d'association est PGCD. Nous noterons aussi, toujours par abus que $\delta = \text{pgcd}(S)$.

Remarque 4.52.
La classe d'association d'un élément n'est pas toujours très grande. Les inversibles dans $\mathbb{Z}$ étant seulement ± 1, nous pouvons obtenir l'unicité du PGCD et du PPCM en imposant qu'ils soient positifs.

Pour les polynômes, nous obtenons l'unicité en demandant que le PGCD soit unitaire.

Dans les cas pratiques, il y a donc en réalité peu d'ambiguïté à parler du PGCD ou du PPCM d'un ensemble.

4.6.4 Anneaux intègres et corps

Le fait d'être intègre pour un anneau n'assure pas le fait d'être un corps. Nous avons cependant ce résultat pour les anneaux finis.

Proposition 4.53.
Un anneau fini intègre est un corps.

Démonstration. Soit A un tel anneau. Soit $a \neq 0$. Les applications

$$l_a : x \to ax \tag{4.35a}$$
$$r_a : x \to xa \tag{4.35b}$$

sont injectives. En tant que applications injectives entre ensembles finis, elles sont surjectives. Il existe donc b et c tels que $1 = ba = ac$. Il se fait que b et c sont égaux parce que

$$b = b(ac) = (ba)c = c. \tag{4.36}$$

Par conséquent b est un inverse de a. □

Il faut être un peu souple avec les notations : ici $b(ac)$ est le produit de b par l'élément bc, et non quelque chose comme le produit de b avec l'idéal (bc).

Proposition 4.54.
Soit $n \in \mathbb{N}^*$. Les conditions suivantes sont équivalentes :

(1) n est premier.

(2) $\mathbb{Z}/n\mathbb{Z}$ est un anneau intègre.

(3) $\mathbb{Z}/n\mathbb{Z}$ est un corps.

Démonstration. L'équivalence entre les deux premiers points est le contenu du corollaire 4.42. Le fait que $\mathbb{Z}/n\mathbb{Z}$ soit un corps lorsque $\mathbb{Z}/n\mathbb{Z}$ est intègre est la proposition 4.53. Le fait que $\mathbb{Z}/n\mathbb{Z}$ soit intègre lorsque $\mathbb{Z}/n\mathbb{Z}$ est un corps est une propriété générale des corps : ce sont en particulier des anneaux intègres (lemme 4.38). □

Proposition 4.55.
Si A est un anneau, nous avons les équivalences

(1) A est un corps.

(2) A est non nul et ses seuls idéaux à gauche sont $\{0\}$ et A.

(3) A est non nul et ses seuls idéaux à droite sont $\{0\}$ et A.

Proposition 4.56.
Soit A, un anneau commutatif non nul et I, un idéal dans A. L'ensemble I est un idéal maximal de A si et seulement si A/I est un corps.

Démonstration. Par la proposition 4.55, le fait pour A/I d'être un corps signifie qu'il n'a pas d'idéaux non triviaux. Si ϕ est la projection canonique $A \mapsto A/I$, nous savons que les idéaux de A/I sont les $\phi(J)$ où J est un idéal de A contenant I (proposition 4.15). Si I est un idéal maximal de A, un tel J n'existe pas. Inversement si A/I n'a pas d'autres idéaux que A/I et $\phi(0)$, c'est que I est un idéal maximal. □

4.6.5 Corps des fractions

Théorème-définition 4.57.
Soit $\mathbb{A}$ un anneau commutatif intègre.

(1) Il existe un corps commutatif $\mathbb{K}$ et un morphisme d'anneaux injectif $\epsilon\colon \mathbb{A} \to \mathbb{K}$ tels que pour tout $\lambda \in \mathbb{K}$, il existe $(a, b) \in \mathbb{A} \times \mathbb{A}^*$ tels que

$$\lambda = \epsilon(a)\big(\epsilon(b)\big)^{-1} \tag{4.37}$$

(2) Si $(\mathbb{K}', \epsilon')$ est un autre couple qui vérifie la propriété, les corps $\mathbb{K}$ et $\mathbb{K}'$ sont isomorphes.

*Le corps $\mathbb{K}$ associé à l'anneau $\mathbb{A}$ est le **corps des fractions** de $\mathbb{A}$, et sera noté* $\mathrm{Frac}(\mathbb{A})$.

Lemme 4.58 (Simplification de fraction)**.**
L'application $\mathbb{A} \times \mathbb{A}^* \to \mathbb{K}$ donnée par $(a, b) \mapsto \epsilon(a)\big(\epsilon(b)\big)^{-1}$ envoie (xa, xb) sur le même que (a, b).

Proposition 4.59.
À propos des rationnels.

(1) L'ensemble $\mathbb{Q}$ est le corps des fractions de $\mathbb{Z}$.

(2) Cet ensemble n'est pas l'ensemble de tous les réels.

4.7 Anneau factoriel

Définition 4.60 (Élément irréductible[31])**.**
*Un élément d'un anneau commutatif intègre est **irréductible** si il n'est ni inversible, ni le produit de deux éléments non inversibles.*

Pour rappel, nous notons $U(A)$ l'ensemble des éléments inversibles de A.

Exemple 4.61
Un corps n'a pas d'éléments irréductibles parce qu'à part zéro tous les éléments sont inversibles. Mais 0 n'est pas irréductible parce qu'il peut être écrit comme produit d'éléments non inversibles : $0 = 0 \cdot 0$. △

Exemple 4.62
Les éléments irréductibles de l'anneau $\mathbb{Z}$ sont les nombres premiers. En effet les seuls inversibles de $\mathbb{Z}$ sont ± 1. Si p est premier et $p = ab$ avec $a, b \in \mathbb{Z}$, alors nous avons soit $a = \pm 1$ soit $b = \pm 1$. △

Définition 4.63 (Anneau factoriel)**.**
*Un anneau commutatif A est **factoriel** s'il vérifie les propriétés suivantes.*

(1) L'anneau A est intègre (pas de diviseurs de zéro).

(2) Si $a \in A$ est non nul et non inversible, alors il admet une décomposition en facteurs irréductibles : $a = p_1 \dots p_k$ où les p_i sont irréductibles.

(3) Si $a = q_1 \dots q_m$ est une autre décomposition de a en irréductibles, alors $m = k$ et il existe une permutation[8] $\sigma \in S_k$ telle que p_i et $q_{\sigma(i)}$ soient associés.

Un anneau factoriel permet de caractériser le pgcd et le ppcm de nombres.

Proposition 4.64.
Soit une famille $\{a_n\}$ d'éléments de A qui se décomposent en irréductibles comme

$$a_i = \prod_k p_k^{\alpha_{k,i}}. \tag{4.38}$$

Alors

$$\text{pgcd}\{a_n\} = \prod_k p_k^{\min_i\{\alpha_{k,i}\}}. \tag{4.39}$$

De plus le PGCD est :

(1) Un multiple de tous les diviseurs communs des a_i.

(2) Unique pour cette propriété à multiple près par un inversible[9] *.*

De la même manière,

$$\text{ppcm}\{a_n\} = \prod_k p_k^{\max_i\{\alpha_{k,i}\}}. \tag{4.40}$$

Un anneau factoriel a une relation de préordre partiel donnée par $a < b$ si a divise b. En termes d'idéaux, cela donne l'ordre inverse de celui de l'inclusion[10] : $a < b$ si et seulement si $(b) \subset (a)$.

Exemple 4.65
L'anneau $\mathbb{Z}[i\sqrt{3}]$ n'est pas factoriel parce que

$$4 = 2 \cdot 2 = (1 + i\sqrt{3})(1 - i\sqrt{3}) \tag{4.41}$$

donnent deux décompositions distinctes de 4 en irréductibles. △

Nous allons voir dans l'exemple 4.86 que $\mathbb{Z}[i\sqrt{2}]$ est factoriel parce qu'il sera euclidien.

4.8 Anneau principal

Définition 4.66.
*Un idéal I dans A est **principal à gauche** s'il existe $a \in I$ tel que $I = Aa$. Il est **principal à droite** s'il existe $a \in I$ tel que $I = aA$. Nous disons qu'il est **principal** s'il est principal à gauche et à droite.*

*Un idéal I dans l'anneau A est **maximal** si les seuls idéaux de A contenant I sont I et A.*

8. Définition 3.90.
9. Soyez prudent avec cette affirmation : je n'en n'ai pas de démonstrations sous la main et ne suis pas certain que ce soit vrai.
10. Voir proposition 4.49.

Définition 4.67.
Un anneau est ***principal*** *si*

(1) il est commutatif et intègre

(2) tous ses idéaux sont principaux.

Souvent pour prouver qu'un anneau est principal, nous prouvons qu'il est euclidien (définition 4.82) et nous utilisons la proposition 4.84 qui dit qu'un anneau euclidien est principal.

Définition 4.68.
Nous disons qu'un idéal I dans A est ***premier*** *si I est strictement inclus dans A et si pour tout $a, b \in A$ tels que $ab \in I$ nous avons $a \in I$ ou $b \in I$.*

Proposition 4.69.
Si A est un anneau commutatif intègre, alors un idéal I dans A est premier si et seulement si A/I est intègre.

Proposition 4.70.
Soit A un anneau principal qui n'est pas un corps. Pour un idéal $I \subset A$, les conditions suivantes sont équivalentes :

(1) I est un idéal maximal ;

(2) I est un idéal premier non nul ;

(3) il existe p irréductible dans A tel que $I = (p)$.

Proposition 4.71.
Si A est un anneau principal et si p est irréductible, alors $A/(p)$ est un corps.

Exemple 4.72
L'anneau $\mathbb{Z}$ est principal parce que ses seuls idéaux sont les $n\mathbb{Z}$ qui sont principaux : $n\mathbb{Z}$ est engendré par n. △

Exemple 4.73
Les anneaux $\mathbb{Z}/n\mathbb{Z}$ sont principaux. En effet les idéaux de $\mathbb{Z}/n\mathbb{Z}$ seraient par la proposition 4.15 des quotients d'idéaux de $\mathbb{Z}$ par (n). Donc les idéaux de $\mathbb{Z}/n\mathbb{Z}$ sont isomorphes aux anneaux $\mathbb{Z}/m\mathbb{Z}$ avec m divisant n.

Par exemple si $I = 4\mathbb{Z}$, on peut considérer l'idéal $J = 2\mathbb{Z}$ qui contient $4\mathbb{Z}$. Donc $\mathbb{Z}/2\mathbb{Z}$ est isomorphe à un idéal de $\mathbb{Z}/4\mathbb{Z}$. △

Définition 4.74.
Nous disons que deux éléments d'un anneau principal sont ***premiers entre eux*** *si leur PGCD est* 1.

Théorème 4.75.
Si A est un anneau principal et si p et q sont premiers entre eux dans A, alors on a l'isomorphisme d'anneaux

$$A/pqA \simeq A/pA \times A/qA. \tag{4.42}$$

4.8.1 Bézout

Théorème 4.76 ([30])**.**
Toute partie S d'un anneau principal admet un PGCD et un PPCM. De plus

$$\delta = \text{pgcd}(S) \Leftrightarrow (\delta) = \sum_{s\in S}(s)\mu = \text{ppcm}(S) \Leftrightarrow (\mu) = \bigcap_{s\in S}(s) \tag{4.43}$$

Démonstration. Vu que l'anneau A est principal, tous ses idéaux sont principaux et donc engendrés par un seul élément. En particulier il existe $\delta, \mu \in A$ tels que

$$(\delta) = \sum_{s\in S}(s) \tag{4.44a}$$

$$(\mu) = \bigcap_{s\in S}(s) \tag{4.44b}$$

PGCD Montrons ce que δ est un PGCD de S. Pour tout $x \in S$, nous avons $(x) \subset (\delta)$, et donc $\delta \mid x$. Par ailleurs si $d \mid x$ pour tout $x \in S$, nous avons $(x) \subset (d)$ et donc

$$\sum_{x\in S}(x) \subset (d), \tag{4.45}$$

puis $(\delta) \subset (d)$ et finalement $d \mid \delta$.

PPCM Si $x \in S$ nous avons $(\mu) \subset (x)$ et donc $x \mid \mu$. D'autre part si $x \mid m$ pour tout $x \in S$, alors $(m) \subset (x)$ et donc $(m) \subset (\mu)$, finalement $\mu \mid m$.

□

Corollaire 4.77 (Théorème de Bézout[30])**.**
Soit A un anneau principal. Deux éléments $a, b \in A$ sont premiers entre eux si et seulement s'il existe un couple $u, v \in A$ tel que

$$ua + vb = 1. \tag{4.46}$$

À la place de 1 on aurait pu écrire n'importe quel inversible.

Démonstration. Pour cette preuve, nous allons écrire $\operatorname{pgcd}(a, b)$ l'ensemble de PGCD de a et b, c'est à dire la classe d'association d'un PGCD.

Si a et b sont premiers entre eux, alors

$$1 \in \operatorname{pgcd}(a, b) = \sum_{x=a,b}(x) = (a) + (b). \tag{4.47}$$

À l'inverse, si nous avons $ua + vb = 1$, alors $1 \in (a) + (b)$, mais vu que $(a) + (b)$ est un idéal principal, $(1) = (a) + (b)$ et donc $1 \in \operatorname{pgcd}(a, b)$. □

Le lemme de Gauss est une application immédiate de Bézout. Il y aura aussi un lemme de Gauss à propos de polynômes (lemme 5.45), et une généralisation directe au théorème de Gauss, théorème 5.44.

Lemme 4.78 (lemme de Gauss)**.**
Soit A un anneau principal et $a, b, c \in A$ tels que a divise bc. Si a est premier avec c, alors a divise b.

Démonstration. Vu que a est premier avec c, nous avons $\operatorname{pgcd}(a, c) = 1$ et Bézout (3.32) nous donne donc $s, t \in \mathbb{A}$ tels que $sa + tc = 1$. En multipliant par b,

$$sab + tbc = b. \tag{4.48}$$

Mais les deux termes du membre de gauche sont multiples de a parce que a divise bc. Par conséquent b est somme de deux multiples de a et donc est multiple de a. □

Un cas usuel d'utilisation est le cas de $A = \mathbb{N}^*$.

4.8.2 Anneau noetherien

Définition 4.79.
Un anneau est dit ***noetherien*** *si toute suite croissante d'idéaux est stationnaire (à partir d'un certain rang).*

Montrer que tout anneau principal est noetherien est le premier pas pour montrer que tout anneau principal est factoriel.

Lemme 4.80.
Tout anneau principal est noetherien.

Démonstration. Soit (J_n) une suite croissante d'idéaux et J la réunion. L'ensemble J est encore un idéal parce que les J_i sont emboités. Étant donné que l'idéal est principal nous pouvons prendre $a \in J$ tel que $J = (a)$. Il existe N tel que $a \in J_N$. Alors pour tout $n \geqslant N$ nous avons

$$J \subset J_N \subset J_n \subset J. \tag{4.49}$$

La première inclusion est le fait que $J = (a)$ et $a \in J_N$. La seconde est la croissance des idéaux et la troisième est le fait que J est une union. Par conséquent pour tout $n \geqslant N$ nous avons $J_N = J_n = J$. La suite est par conséquent stationnaire. □

Théorème 4.81 ([32])**.**
Tout anneau principal est factoriel.

4.9 Anneau euclidien

Définition 4.82 (Wikipédia)**.**
Soit A un anneau intègre. Un ***stathme euclidien*** *sur A est une application $\alpha\colon A\backslash\{0\} \to \mathbb{N}$ tel que*

(1) $\forall a, b \in A\backslash\{0\}$, il existe $q, r \in A$ tel que

$$a = bq + r \tag{4.50}$$

et $\alpha(r) < \alpha(b)$.

(2) Pour tout $a, b \in A\backslash\{0\}$, $\alpha(b) \leqslant \alpha(ab)$.

Un anneau est ***euclidien*** *s'il accepte un stathme euclidien.*

Le stathme est la fonction qui donne le «degré» à utiliser dans la division euclidienne. La contrainte est que le degré du reste soit plus petit que le degré du dividende.

Exemple 4.83
Le stathme de $\mathbb{N}$ pour la division euclidienne usuelle est $\alpha(n) = n$. Si $a, b \in \mathbb{N}$ nous écrivons

$$a = bq + r \tag{4.51}$$

où q est l'entier le plus proche *inférieur* à a/b (on veut que le reste soit positif) et $r = a - bq$. Nous avons donc

$$r - b = a - b(q + 1) < a - b\frac{a}{b} = 0, \tag{4.52}$$

ce qui montre que $r < b$. △

Cet exemple ne fonctionne pas avec $\mathbb{Z}$ au lieu de $\mathbb{N}$ parce que le stathme doit avoir des valeurs dans $\mathbb{N}$. Cela ne veut cependant pas dire qu'il n'existe pas de stathme sur $\mathbb{Z}$; cela veut seulement dire que $\alpha(x) = x$ ne fonctionne pas. L'exemple 4.85 montrera que $\mathbb{Z}$ ne possède pas de stathme euclidien en montrant qu'il n'est pas principal.

Proposition 4.84 ([33])**.**
Un anneau euclidien est principal.

Démonstration. Soit A un anneau principal et α un stathme sur A. Nous considérons un idéal I non nul de A. Nous devons montrer que I est généré par un élément. En l'occurrence nous allons montrer qu'un élément $a \in I\backslash\{0\}$ qui minimise $\alpha(a)$ va générer [11]. Soit $x \in I$. Par construction, il existe $q, r \in A$ tels que $x = aq + r$ avec $r = 0$ ou $\alpha(r) < \alpha(a)$. Étant donné que $x, a \in I$, $r \in I$. Si $r \neq 0$, alors r contredirait la minimalité de $\alpha(a)$. Donc $r = 0$ et $x = aq$, ce qui signifie que I est principal. □

Exemple 4.85($\mathbb{Z}$ n'est ni principal ni euclidien)
L'idéal $(2) = 2\mathbb{Z}$ n'est pas un idéal principal de $\mathbb{Z}$, parce que tout multiple entier d'un nombre pair reste pair. Donc $\mathbb{Z}$ possède des idéaux non principaux et n'est pas conséquent pas un anneau principal.

Il ne peut alors pas non plus être euclidien par la proposition 4.84. △

Exemple 4.86
Prouvons que $\mathbb{Z}[i\sqrt{2}]$ est une anneau euclidien. Pour cela nous démontrons que

$$\begin{aligned} N\colon \mathbb{Z}[i\sqrt{2}] &\to \mathbb{N} \\ a + bi\sqrt{2} &\mapsto a^2 + 2b^2 \end{aligned} \tag{4.53}$$

est un stathme euclidien.

Soient $z = a + bi\sqrt{2}$, $t = a' + b'i\sqrt{2}$. Nous cherchons q et r tels que la division euclidienne s'écrive $z = qt + r$. Soient $\alpha, \beta \in \mathbb{Q}$ tels que

$$\frac{z}{t} = \alpha + \beta i\sqrt{2}. \tag{4.54}$$

Nous désignons par $\alpha + \epsilon_1$ et $\beta + \epsilon_2$ les entiers les plus proches de α et b. Nous avons $|\epsilon_1|, |\epsilon_2| \leqslant \frac{1}{2}$. Nous posons alors naturellement

$$q = (\alpha + \epsilon_1) + (\beta + \epsilon_2)i\sqrt{2} \tag{4.55}$$

et nous calculons $r = z - qt$:

$$2b'\epsilon_2 - a'\epsilon_1 + i\sqrt{2}\big(\epsilon_1 b' - a'\epsilon_2\big). \tag{4.56}$$

Nous trouvons

$$N(r) = a'^2\epsilon_1^2 + 4b'^2\epsilon_2^2 + 2a'^2\epsilon_1^2 + 2b'^2\epsilon_2^2 \leqslant \frac{3}{4}a'^2 + \frac{3}{2}b'^2. \tag{4.57}$$

D'autre part $N(t) = a'^2 + 2b'^2$, et nous avons donc bien $N(r) < N(t)$.

En ce qui concerne la seconde propriété du stathme, un petit calcul montre que

$$N(zt) = (a^2 + 2b^2)(a'^2 + 2b'^2), \tag{4.58}$$

et tant que $t \neq 0$ nous avons bien $N(zt) > N(z)$. △

Notons en particulier que $\mathbb{Z}[i\sqrt{2}]$ est factoriel et principal.

Exemple 4.87
Décomposition en facteurs irréductibles dans $\mathbb{Z}[i\sqrt{2}]$. Les éléments inversibles de $\mathbb{Z}[i\sqrt{2}]$ sont ± 1, donc deux éléments a et b sont associés (définition 4.47) si et seulement si $a = \pm b$.

De plus si p est irréductible, alors $-p$ est irréductible. Les éléments irréductibles de $\mathbb{Z}[i\sqrt{2}]$ arrivent donc par pairs d'éléments associés. Soit $\{p_i\}$ une sélection de un élément irréductible

11. Un tel élément existe...

parmi chaque paire. Tout élément x de $\mathbb{Z}[i\sqrt{2}]$ peut alors être écrit $x = \pm p_1^{\alpha_1} \dots p_n^{\alpha_n}$. Ce fait va être pratique pour comparer des décomposition en facteurs irréductibles d'éléments. △

Le lemme suivant fait en pratique partie de l'exemple 4.90, mais nous l'isolons pour plus de clarté [12].

Lemme 4.88.
Si a et b sont deux éléments premiers entre eux de $\mathbb{Z}[i\sqrt{2}]$, et s'il existe $y \in \mathbb{Z}[i\sqrt{2}$ tel que $ab = y^3$, alors a et b sont des cubes (dans $\mathbb{Z}[i\sqrt{2}]$).

Démonstration. D'après l'exemple 4.87 nous pouvons écrire

$$y = \pm p_1^{\sigma_1} \dots p_n^{\sigma_n} \tag{4.59a}$$
$$a = \pm p_1^{\alpha_1} \dots p_n^{\alpha_n} \tag{4.59b}$$
$$b = \pm p_1^{\beta_1} \dots p_n^{\beta_n} \tag{4.59c}$$

où les p_i sont les irréductibles de $\mathbb{Z}[i\sqrt{2}]$ «modulo ± 1» au sens où la liste des irréductibles est $\{p_i\} \cup \{-p_i\}$ (union disjointe). Étant donné que a et b sont premiers entre eux, α_i et β_i ne peuvent pas être non nuls en même temps alors que leur somme doit faire $3\sigma_i$. Nous avons donc pour chaque i soit $\alpha_i = 3\sigma_i$ soit $\beta = 3\sigma_i$ (et bien entendu si $\sigma_i = 0$ alors $\alpha_i = \beta_i = 0$).

Étant donné que ± 1 sont également deux cubes, a et b sont bien des cubes.

Notons que nous avons utilisé de façon capitale le fait que $\mathbb{Z}[i\sqrt{2}]$ était factoriel. □

4.9.1 Équations diophantiennes

Exemple 4.89
L'équation diophantienne

$$x^2 = 3y^2 + 8 \tag{4.60}$$

n'a pas de solutions. En effet si nous prenons l'équation modulo 3 nous obtenons

$$[x^2]_3 = [3y^2 + 8]_3 = [8]_3 = [2]_3. \tag{4.61}$$

Or dans $\mathbb{Z}/3\mathbb{Z}$, aucun carré n'est égal à deux : $0^2 = 0 \neq 2$, $1^2 = 1 \neq 2$ et $2^2 = 4 = 1 \neq 2$. △

Exemple 4.90
Résolvons l'équation diophantienne

$$x^2 + 2 = y^3. \tag{4.62}$$

Une première remarque est que x doit être impair. En effet si $x = 2k$, nous devons avoir y^3 pair. Mais si un cube pair est divisible par 8, donc $y^3 = 8l$. L'équation devient $4k^2 + 2 = 8l^3$, c'est à dire $2k^2 + 1 = 4l^3$. Le membre de gauche est impair tandis que celui de droite est pair. Impossible.

Nous pouvons écrire l'équation sous la forme $x^2 + 2 = (x + i\sqrt{2})(x - i\sqrt{2})$. Et nous considérons $\mathbb{Z}[i\sqrt{2}]$ muni de son stathme N donné par (4.53).

L'élément $i\sqrt{2}$ est irréductible parce que $N(i\sqrt{2}) = 2$, et si nous avions $i\sqrt{2} = pq$, alors nous aurions $N(p)N(q) = 2$, ce qui n'est possible que si $N(p)$ ou $N(q)$ vaut 1.

Nous prouvons maintenant que les éléments $x + i\sqrt{2}$ et $x - i\sqrt{2}$ sont premiers entre eux. Supposons que d soit un diviseur commun ; alors il divise aussi la somme et la différence. Donc d divise à la fois $2x$ et $2i\sqrt{2}$.

Étant donné que $i\sqrt{2}$ est irréductible et que $2i\sqrt{2} = (-i\sqrt{2})^3$, les diviseurs de $2i\sqrt{2}$ sont les puissances de $(-i\sqrt{2})$. Du coup nous devrions avoir $d = (i\sqrt{2})^\beta$ et donc

$$x = (i\sqrt{2})^\beta q \tag{4.63}$$

12. Merci à Marvoir pour m'avoir souligné le manque.

pour un certain $q \in \mathbb{Z}[i\sqrt{2}]$. Dans ce cas nous avons $N(x) = 2^\beta N(q)$, mais nous avons déjà précisé que x ne pouvait pas être pair, donc $\beta = 0$ et nous avons $d = 1$.

Vu que les nombres $x \pm i\sqrt{2}$ sont premiers entre eux et que leur produit doit être un cube, ils doivent être séparément des cubes (lemme 4.88). Nous devons donc résoudre séparément $x \pm i\sqrt{2} = y^3$.

Cherchons les x et y entiers tels que $x + i\sqrt{2} = y^3$. Si nous posons $z = a + bi\sqrt{2}$, il suffit de calculer z^3 :

```
----------------------------------------------------------------------
| Sage Version 4.8, Release Date: 2012-01-20                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('a,b')
(a, b)
sage: z=a+I*sqrt(2)*b
sage: (z**3).expand()
3*I*sqrt(2)*a^2*b - 2*I*sqrt(2)*b^3 + a^3 - 6*a*b^2
```

En identifiant cela à $x + i\sqrt{2}$ nous trouvons le système

$$\begin{cases} x = a^3 - 6ab^2 & \text{(4.64a)} \\ 1 = 3a^2b - 2b^3 & \text{(4.64b)} \end{cases}$$

où, nous le rappelons, x, a et b sont des entiers. La seconde équation montre que b doit être inversible : $b(3a^2 - 2b^2) = 1$. Il y a donc les possibilités $b = \pm 1$. Pour $b = 1$ l'équation devient $3a^2 - 2 = 1$, c'est à dire $a = \pm 1$. Pour $b = -1$ l'équation devient $3a^2 - 2 = -1$, impossible. En conclusion les possibilités sont

$$(x, z) = (-5, 1 + i\sqrt{2}) \quad \text{(4.65a)}$$
$$(x, z) = (5, -1 + i\sqrt{2}) \quad \text{(4.65b)}$$
$$\text{(4.65c)}$$

Le travail avec $x - i\sqrt{2}$ donne les mêmes résultats.

Les deux solutions de l'équation $x^2 + 2 = y^3$ sont alors $(5, 3)$ et $(-5, 3)$. △

4.9.2 Triplets pythagoriciens et équation de Fermat pour $n = 4$

Définition 4.91.
*Les solutions entières (positives) de l'équation $x^2 + y^2 = z^2$ sont appelés **triplets pythagoriciens**.*

Ils donnent toutes les possibilités de triangles rectangles dont les côtés ont des longueurs entières.

Définition 4.92.
*On dit qu'un triplet pythagoricien est **primitif** si les trois nombres sont premiers dans leur ensemble*[13].

Remarquons que cela est équivalent à montrer que les trois nombres sont premiers deux à deux : en effet, si deux parmi x, y et z sont divisibles par un nombre, alors tous les trois sont divisibles par ce nombre[14], donc les nombres x, y et z sont premiers deux à deux.

Lemme 4.93.
Dans un triplet pythagoricien primitif (x, y, z), on a toujours z impair et :

13. Définition 3.30.
14. Parce que k et k^2 ont les mêmes facteurs premiers.

— *soit x impair et y pair ;*
— *soit x pair et y impair.*

Démonstration. Remarquons que le fait d'imposer que le triplet soit primitif, interdit aux nombres x et y d'être pairs en même temps. En effet, si c'était le cas, alors x^2 et y^2 seraient aussi pairs, donc leur somme z^2 aussi, d'ou z serait pair et les trois nombres ne seraient pas premiers entre eux.

Nous montrons à présent que les nombres x et y ne sont pas tous les deux impairs. Par l'absurde, si $x = 2a + 1$, nous avons $x^2 = 4a^2 + 4a + 1 \in [1]_4$; de la même manière, $y^2 \in [1]_4$. On en déduit alors que $z^2 = x^2 + y^2 \in [2]_4$. Le nombre z^2 est donc pair, donc z est pair : disons $z = 2c$. Alors, $z^2 = 4c^2 \in [0]_4$. Comme les classes modulo 4 sont disjointes, nous aboutissons à une contradiction. □

Proposition 4.94 (Triplets pythagoriciens[34, 35])**.**
Un triplet $(x, y, z) \in (\mathbb{N}^*)^3$ est solution de $x^2+y^2 = z^2$ si et seulement s'il existe $d \in \mathbb{N}$ et $u, v \in \mathbb{N}^*$ premiers entre eux tels que

$$\begin{cases} x = d(u^2 - v^2) & (4.66a) \\ y = 2duv & (4.66b) \\ z = d(u^2 + v^2) & (4.66c) \end{cases}$$

ou

$$\begin{cases} x = 2duv & (4.67a) \\ y = d(u^2 - v^2) & (4.67b) \\ z = d(u^2 + v^2) & (4.67c) \end{cases}$$

La différence entre les deux est seulement d'inverser les rôles de x et y.

Démonstration. Montrons d'abord que les formules proposées sont bien des solutions ; nous vérifions (4.66) :

$$x^2 + y^2 = d^2(u^2 - v^2) + 4d^2u^2v^2 = d^2(u^2 + v^2)^2, \tag{4.68}$$

qui correspond bien au z^2 proposé.

Nous allons maintenant prouver la réciproque : toute solution est d'une des deux formes proposées. Déterminer les triplets primitifs suffira parce que si (x, y, z) n'est pas une solution primitive, alors en posant $k = \text{pgcd}(x, y, z)$, le triplet $\left(\frac{x}{k}, \frac{y}{k}, \frac{z}{k}\right)$ est primitif. Connaissant les triplets primitifs, nous obtenons tous les autres par simple multiplication.

Soit donc (x, y, z) un triplet pythagoricien primitif. Sans perte de généralité [15], grâce au lemme 4.93, on peut supposer x pair, y impair, z impair. Comme $x^2 = (z + y)(z - y)$, nous avons

$$\left(\frac{x}{2}\right)^2 = \left(\frac{z+y}{2}\right)^2 \left(\frac{z-y}{2}\right)^2. \tag{4.69}$$

Vu que z et y sont premiers entre eux, les nombres $z - y$ et $z + y$ sont également premiers entre eux [16]. Donc les facteurs premiers (qui arrivent au moins au carré) de $(x/2)^2$ sont chacun soit dans $(z + y)/2$ soit dans $(z - y)/2$. Nous en déduisons que ces derniers sont des carrés d'entiers. Nous posons

$$\frac{z-y}{2} = u^2 \quad \frac{z+y}{2} = v^2. \tag{4.70}$$

Bien entendu u et v sont non nuls parce que nous avons exclu la possibilité de triplets dont un élément serait nul. Avec tout cela nous avons $(x/2)^2 = u^2v^2$ et donc $x = 2uv$ puis par somme et différence :

$$\begin{cases} x = 2uv & (4.71a) \\ y = v^2 - u^2 & (4.71b) \\ z = u^2 + v^2, & (4.71c) \end{cases}$$

15. En échangeant les rôles de x et y ici, nous obtenons à la fin la seconde forme des solutions.
16. Si $z - y = kn$ et $z + y = km$, faisant la somme et la différence on voit que y et z sont divisibles par k.

ce qu'il fallait. □

Remarque 4.95.
Les solutions dans lesquelles x, y ou z sont nuls sont faciles à classer. La solution $(1,0,1)$ n'est pas dans les formes proposées. En effet elle ne peut pas être de la première forme : avoir $y=0$ demanderait qu'un nombre parmi d, u et v soit nul, ce qui est interdit. La solution $(1,0,1)$ ne peut pas non plus être de la seconde forme parce que x y est pair.

Proposition 4.96 ([34])**.**
Les équations $x^4+y^4=z^2$ et $x^2+y^4=z^4$ n'ont pas de solutions dans $(\mathbb{N}^*)^3$.

Démonstration. Si la première équation n'a pas de solutions, alors la seconde n'en n'a pas non plus parce que z^4 est un carré. Nous nous concentrons donc sur l'équation $x^4+y^4=z^2$ et nous supposons qu'il existe au moins une solution dans $(\mathbb{N}^*)^3$. Nous en choisissons une (x,y,z) avec z minimum (les z dans différentes solutions étant dans $\mathbb{N}$, il en existe forcément un minimum [17]). Du coup, les trois nombres x, y et z sont premiers dans leur ensembles parce que une division par leur pgcd donnerait une nouvelle solution qui contredirait la minimalité de z.

Nous posons $x^4=\bar{x}^2$ et $y^4=\bar{y}^2$. Ils vérifient l'équation $\bar{x}^2+\bar{y}^2=z^2$ et par la proposition 4.94, il existe $u,v\in\mathbb{N}^*$ premiers entre eux tels que, sans perte de généralité [18], on ait

$$\begin{cases} \bar{x}=2uv & \text{(4.72a)}\\ \bar{y}=u^2-v^2 & \text{(4.72b)}\\ z=u^2+v^2. & \text{(4.72c)}\end{cases}$$

Si u est pair, alors v est impair (et inversement) parce que $\operatorname{pgcd}(u,v)=1$ Si u est pair, alors $u=2a$ et $v=2b+1$, ce qui donne $\bar{y}=4a^2-4b^2-4b-1\in[-1]_4$. Or nous avons déjà vu qu'un carré est dans $[0]_4$ ou dans $[1]_4$. Il faut donc que u soit impair. Le lemme 4.93 implique alors que v soit pair.

De l'équation 4.72b, nous en déduisons que $v^2+\bar{y}=u^2$; de plus u^2, v^2 et $\bar{y}$ sont premiers dans leur ensemble : en effet, u et v sont premiers entre eux, et si l'un parmi u^2 et v^2 a un facteur commun avec $\bar{y}$, alors l'autre doit l'avoir aussi (dans une égalité $a+b=c$, si deux des nombres ont un diviseur commun, le troisième l'a aussi). Comme $\bar{y}=y^2$, le triplet (v,y,u) est un triplet pythagoricien primitif. Nous réappliquons la proposition 4.94, en se souvenant que v est pair : il existe donc deux nombres r et s premiers entre eux tels que

$$\begin{cases} v=2rs & \text{(4.73a)}\\ y=r^2-s^2 & \text{(4.73b)}\\ u=r^2+s^2. & \text{(4.73c)}\end{cases}$$

Avec cela, $\bar{x}=2uv=4rs(r^2+s^2)$. Remarquons que les trois nombres r, s et r^2+s^2 sont premiers entre eux dans leur ensemble ; or, comme $\bar{x}$ est un carré ces nombres doivent séparément être des carrés :

$$\begin{cases} r=\alpha^2 & \text{(4.74a)}\\ s=\beta^2 & \text{(4.74b)}\\ r^2+s^2=\gamma^2. & \text{(4.74c)}\end{cases}$$

En mettant les deux premiers dans la troisième, nous avons $\alpha^4+\beta^4=\gamma^2$. Donc $(\alpha^2,\beta^2,\gamma)$ est une solution. Nous allons prouver que $\gamma<z$, ce qui terminera la preuve, puisque z était supposé

17. Voir quelque chose comme le lemme 2.17.
18. En inversant les rôles de x et y au besoin.

minimal. Nous avons :

$$\begin{aligned} z &= u^2 + v^2 && \text{par 4.72c} \\ &= r^2 + s^2 + 4r^2s^2 && \text{par 4.73} \\ &= \gamma^2 + 4r^2s^2 \\ &> \gamma^2, \end{aligned}$$

et a fortiori $\gamma < z$. □

4.9.3 Lignes et colonnes de matrices

Nous nommons E_{ij} la matrice remplie de zéros sauf à la case ij qui vaut 1. Autrement dit

$$(E_{ij})_{kl} = \delta_{ik}\delta_{jl}. \tag{4.75}$$

Définition 4.97.
*Une **matrice de transvection** est une matrice de la forme*

$$T_{ij}(\lambda) = \mathrm{Id} + \lambda E_{ij} \tag{4.76}$$

avec $i \neq j$.

*Une **matrice de dilatation** est une matrice de la forme*

$$D_i(\lambda) = \mathrm{Id} + (\lambda - 1)E_{ii}. \tag{4.77}$$

Ici le $(\lambda - 1)$ sert à avoir λ et non $1 + \lambda$. C'est donc une matrice qui dilate d'un facteur λ la direction i tout en laissant le reste inchangé.

*Si σ est une permutation (un élément du groupe symétrique S_n) alors la **matrice de permutations** associée est la matrice d'entrées*

$$(P_\sigma)_{ij} = \delta_{i\sigma(j)}. \tag{4.78}$$

Lemme 4.98.
La matrice $T_{ij}(\lambda)A = (\mathbb{1} + \lambda E_{ij})A$ est la matrice A à qui on a effectué la substitution

$$L_i \to L_i + \lambda L_j. \tag{4.79}$$

La matrice $AT_{ij}(\lambda)$ est la substitution

$$C_j \to C_j + \lambda C_i. \tag{4.80}$$

La matrice AP_σ est la matrice A dans laquelle nous avons permuté les colonnes avec σ.
La matrice $P_\sigma A$ est la matrice A dans laquelle nous avons permuté les lignes avec σ^{-1}.

Démonstration. Calculons la composante kl de la matrice $E_{ij}A$:

$$\begin{aligned} (E_{ij}A)_{kl} &= \sum_m (E_{ij})_{km}A_{ml} && (4.81a) \\ &= \sum_m \delta_{ik}\delta_{jm}A_{ml} && (4.81b) \\ &= \delta_{ik}A_{jl}. && (4.81c) \end{aligned}$$

C'est donc la matrice pleine de zéros, sauf la ligne i qui est donnée par la ligne j de A. Donc effectivement la matrice

$$A + \lambda E_{ij}A \tag{4.82}$$

est la matrice A à laquelle on a substitué la ligne i par la ligne i plus λ fois la ligne j.

En ce qui concerne l'autre assertion sur les transvections, le calcul est le même et nous obtenons

$$(AE_{ij}) = A_{ki}\delta_{jl}. \tag{4.83}$$

Pour les matrices de permutations, nous avons

$$(AP_\sigma)_{kl} = A_{k\sigma(l)} \tag{4.84}$$

et

$$(P_\sigma A)_{kl} = \sum_m \delta_{k\sigma(m)}A_{ml} = \sum_m \delta_{\sigma^{-1}(k)m}A_{ml} = A_{\sigma^{-1}(k)l}. \tag{4.85}$$

□

4.9.4 Algorithme des facteurs invariants

Problèmes et choses à faire

Définir les anneaux de matrices, et en particulier $GL_n(A)$.

Proposition 4.99 (Algorithme des facteurs invariants[2])**.**
Soit (A, δ) un anneau euclidien muni de son stathme et $U \in \mathbb{M}(n, m, A)$. Alors il existe $d_1, \dots, d_s \in A^*$ et des matrices $P \in \mathrm{GL}(m, A)$, $Q \in \mathrm{GL}(n, A)$ tels que nous ayons

$$U = P \begin{pmatrix} d_1 & & & \\ & \ddots & & 0 \\ & & d_s & \\ & 0 & & 0 \end{pmatrix} Q \tag{4.86}$$

avec $d_i \mid d_{i+1}$ pour tout i.

Démonstration. Nous allons donner la preuve plus ou moins sous forme d'algorithme.

D'abord si $U = 0$ c'est bon, on a la réponse. Sinon, nous prenons l'élément (i_0, j_0) dont le stathme est le plus petit et nous l'amenons en $(1, 1)$ par les permutations

$$\begin{aligned} C_1 &\leftrightarrow C_{j_0} \\ L_1 &\leftrightarrow L_{i_0} \end{aligned} \tag{4.87}$$

Ensuite nous traitons la première colonne jusqu'à amener des zéros partout en dessous de u_{11} de la façon suivante : pour chaque ligne successivement nous calculons la division euclidienne

$$u_{i1} = qu_{11} + r_i, \tag{4.88}$$

et nous faisons

$$L_i \to L_i - qL_1, \tag{4.89}$$

c'est à dire que nous enlevons le maximum possible et il reste seulement r_i en u_{i1}. Vu que le but est de ne laisser que des zéros dans la première colonne, si le reste n'est pas zéro nous ne sommes pas content [19]. Dans ce cas nous permutons $L_1 \leftrightarrow L_i$, ce qui aura pour effet de strictement diminuer le stathme de u_{11} parce qu'on va mettre en u_{11} le nombre r_i dont le stathme est strictement plus petit que celui de u_{11}.

En faisant ce jeu de division euclidienne puis échange, on diminue toujours le stathme de u_{11}, donc ça finit par s'arrêter, c'est à dire qu'à un certain moment la division euclidienne de u_{i1} par u_{11} va donner un reste zéro et nous serons content.

19. S'il est zéro, nous passons à la ligne suivante

Une fois la première colonne ramenée à la forme

$$C_1 = \begin{pmatrix} u_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{4.90}$$

nous faisons tout le même jeu avec la première ligne en faisant maintenant des sommes divisions et permutations de colonnes. Notons que ce faisant nous ne changeons plus la première colonne.

En fin de compte nous trouvons une matrice [20]

$$U = \begin{pmatrix} u_{11} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix} \tag{4.91}$$

Si l'élément u_{11} ne divise pas un des éléments de A, disons a_{ij}, alors nous faisons

$$C_1 \to C_1 - C_j. \tag{4.92}$$

Cela nous détruit un peu la première colonne, mais ne change pas u_{11}. Nous avons maintnant

$$U = \begin{pmatrix} u_{11} & 0 & \dots & 0 \\ 0 & & & \\ * & & & \\ u_{ij} & & A & \\ * & & & \\ 0 & & & \end{pmatrix} \tag{4.93}$$

Et nous refaisons tout le jeu depuis le début. Cependant lorsque nous allons nous attaquer à la ligne i, u_{11} ne divisera pas u_{ij}, ce qui donnera lieu à une division euclidienne et un échange $L_1 \leftrightarrow L_i$. L'échange consistant à mettre r_i à la place de u_{11} et inversement diminuera encore strictement le stathme. Encore une fois nous allons travailler jusqu'à avoir la matrice sous la forme

$$U = \begin{pmatrix} u_{11} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix}, \tag{4.94}$$

sauf que cette fois le stathme de u_{11} est strictement plus petit que la fois précédente. Si u_{11} ne divise toujours pas tous les éléments de A, nous recommençons encore et encore. En fin de compte nous finissons par avoir une matrice de la forme (4.94) avec u_{11} qui divise tous les éléments de A.

Une fois que cela est fait, il faut continuer en recommençant tout sur la matrice A. Nous avons maintenant

$$U = \begin{pmatrix} u_{11} & & 0 \\ & u_{22} & \\ 0 & & B \end{pmatrix}. \tag{4.95}$$

Sous cette forme nous avons $u_{11} \mid u_{22}$ et u_{11} divise tous les éléments de B. En effet u_{11} divisant tous les éléments de A, il divise toutes les combinaisons de ces éléments. Or tout l'algorithme ne consiste qu'à prendre des combinaisons d'éléments.

Nous finissons donc bien sûr une matrice comme annoncée. De plus n'ayant effectué que des combinaisons de lignes, nous avons seulement multiplié par des matrices inversibles (lemme 4.98). □

20. Nous nommons toujours par la même lettre U la matrice originale et la modifiée, comme il est d'usage en informatique.

4.10 Polynômes à coefficients dans un anneau commutatif

Soit A un anneau commutatif. Nous considérons $\mathcal{P}$ l'ensemble des suites presque nulles d'éléments de A, ce sont les suites $(a_n)_{n\in\mathbb{N}}$ telles que il existe N tel que $a_i = 0$ pour tout $i > N$.

Cela est un A-module libre de base [21]

$$(e_n)_k = \delta_{nk}. \tag{4.96}$$

Si $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont des éléments de $\mathcal{P}$, nous définissons le produit ab par

$$(ab)_n = \sum_{p+q=n} a_p b_q. \tag{4.97}$$

Cela est bien un élément de $\mathcal{P}$ parce qu'il existe $N \in \mathbb{N}$ tel que $a_n = b_n = 0$ pour tout $n \geqslant N$. Avec la somme et le produit par un scalaire, le module $\mathcal{P}$ devient une A-algèbre commutative unitaire. L'unité est

$$e_0 = (1, 0, \ldots). \tag{4.98}$$

Définition 4.100.
*En tant que A-algèbre, l'ensemble $\mathcal{P}$ est l'**algèbre des polynômes en une indéterminée** à coefficients dans A.*

Définition 4.101.
*Soit $P \in \mathcal{P}$, $P \neq 0$. On appelle **degré** de P le plus grand nombre naturel n pour lequel le coefficient correspondant est non-nul. Ce naturel est noté* $\deg(P)$.

Si nous posons que $X = e_1$, et que nous prenons la convention $X^0 = 1$, alors nous avons $e_k = X^k$ et nous notons $A[X]$ l'anneau $\mathcal{P}$ exprimé avec X. Les éléments de la forme λX^k avec $\lambda \in A$ et $k \in \mathbb{N}$ sont des **monômes**. Nous allons aussi considérer

$$A_n[X] = \{P \in A[X] \text{ tel que } \deg(P) \leqslant n\}. \tag{4.99}$$

Cela est un sous module libre.

Remarque 4.102.
L'ensemble $A[X]$ est une algèbre et donc un espace vectoriel. Il possède un unique élément nul qui est celui dont tous les coefficients sont nuls ; cela est immédiat par la construction en tant que suites presque nulles.

Il n'y a a priori pas équivalence entre le fait d'être un polynôme nul et le fait de s'évaluer à zéro sur tous les éléments de A. Cela sera discuté dans le théorème 5.68 et l'exemple 5.134.

Théorème 4.103.
L'anneau A est intègre si et seulement si $A[X]$ est intègre.

Démonstration. Soient P et Q des éléments non nuls de $A[X]$. Vu que l'anneau A est intègre, nous avons

$$\deg(PQ) = \deg(P) + \deg(Q) \tag{4.100}$$

et le produit ne peut pas être nul. L'anneau $A[X]$ est donc intègre.

Si $A[X]$ est intègre, A est intègre parce qu'il peut être vu comme sous anneau. □

4.104.
Si A n'est pas intègre, soient $\alpha, \beta \in A$ non nuls tels que $\alpha\beta = 0$. Le produit est polynômes $X \mapsto \alpha X$ et $X \mapsto \beta$ est $(\alpha X)(\beta) = 0$; le degré du produit n'est pas la somme des degrés.

Les personnes qui ont tout compris jusqu'ici remarqueront que la notation «$X \mapsto P(X)$» n'est pas correcte parce que du point de vue que nous adoptons ici, un polynôme n'est pas une application.

21. Définition 4.27.

Corollaire 4.105.
Si A est intègre, les inversibles de $A[X]$ sont les éléments de $U(A)$.

Démonstration. Pour que Q soit inversible, il faut un P tel que $PQ = 1$. Mais l'anneau A étant intègre, les degrés s'additionnent. Par conséquent ils doivent être de degrés zéro et il faut que $P, Q \in A$. Enfin pour qu'ils soient inversibles, ils doivent être dans $U(A)$. □

La **valuation** du polynôme $P = \sum_n a_n X^n$, notée $\text{val}(P)$, est

$$\text{val}(P) = \min\{n \text{ tel que } a_n \neq 0\}. \tag{4.101}$$

Nous avons $\text{val}(P) \leqslant \deg(P)$ et $\text{val}(P) = \deg(P)$ si et seulement si P est un monôme. Si $P = 0$, nous convenons que $\text{val}(0) = \infty$ et $\deg(0) = -\infty$.

4.10.1 Division euclidienne

Le théorème suivant établit la **division euclidienne** dans $A[X]$ du polynôme P par un polynôme D.

Théorème 4.106.
Soit $D \neq 0$ dans $A[X]$ de coefficient dominant inversible dans A. Pour tout $P \in A[X]$, il existe $Q, R \in A[X]$ tels que

$$P = QD + R \tag{4.102}$$

avec $\deg(R) < \deg(D)$.

Les polynômes Q et R sont déterminés de façon univoque par cette condition.

Définition 4.107.
*Le polynôme Q est le **quotient** et R est le **reste** de la division euclidienne de P par D. Si le reste de la division de P par D est nul on dit que D **divise** P et on note $D \mid P$. Autrement dit D divise P s'il existe Q tel que $P = QD$.* [22]

4.108.
Le théorème 4.106 nous incite à utiliser le degré comme stathme euclidien sur $A[X]$ dès que A est un anneau intègre. Or cela ne fonctionne en général pas parce que très peu de polynômes ont a priori un coefficient dominant inversible.

En exemple typique, l'anneau $\mathbb{Z}[X]$ n'est pas euclidien. Il suffit de reprendre le même argument que celui donné dans l'exemple 4.85 : l'idéal $2\mathbb{Z}[X]$ dans $\mathbb{Z}[X]$ n'est pas principal. Donc $\mathbb{Z}[X]$ n'est pas un idéal principal et par conséquent pas euclidien non plus.

Lemme 4.109.
Si $\mathbb{K}$ est un corps [23]*, alors l'anneau $\mathbb{K}[X]$ est euclidien et principal.*

Démonstration. Vu que $\mathbb{K}$ est un corps, tous les éléments sont inversibles et le degré donne un stathme par le théorème 4.106. L'anneau $\mathbb{K}[X]$ est donc euclidien et par conséquent principal (proposition 4.84). □

Dans le théorème 5.37 nous donnerons une preuve directe du fait que $\mathbb{K}[X]$ est principal en montrant que tous ses idéaux sont principaux. Nous y démontrerons donc un peu moins pour un peu plus cher, mais avec le plaisir de ne pas devoir passer par un stathme.

Définition 4.110 ([36])**.**
*Soit un anneau A. Deux polynômes P et Q dans $A[X]$ sont dits **étrangers** entre eux si* 1 *est un pgcd* [24] *de P et Q. Un ensemble de polynômes $(P_i)_{i\in I}$ est étranger **dans leur ensemble** si* 1 *est un* pgcd *des P_i.*

22. Ceci se rapproche tout naturellement des notions de divisibilité dans un anneau intègre général, vues en sous-section 4.6.2.
23. Définition 2.21.
24. Définition 4.8.

Les polynômes P et Q sont ***premiers entre eux*** *si les seuls diviseurs communs de P et Q sont les inversibles.*

Les notions de polynômes étrangers entre eux ou de polynômes premiers entre eux ne sont pas identiques, comme le montre l'exemple suivant.

Exemple 4.111([1])
Soient dans $\mathbb{Z}[X]$ les polynômes $P(X) = 2X + 2$ et $Q(X) = 2X^2 + 2$. Le nombre 2 est diviseur commun et n'est pas un diviseur de 1. Donc 1 n'est pas un pgcd de P et Q. Ils ne sont pas étrangers.

Mais ils sont premiers entre eux parce qu'ils n'ont pas d'autres diviseurs communs que les inversibles (1 et -1). △

4.10.2 Polynôme primitif

Définition 4.112.
Le ***contenu*** *du polynôme $P = \sum_i a_i X^i \in \mathbb{K}[X]$ est le pgcd de ses coefficients : $c(P) = \text{pgcd}(a_i)$.*

Définition 4.113 (Ordre d'un polynôme)**.**
*Soit P un polynôme irréductible de degré n sur $\mathbb{F}_p[X]$. L'****ordre*** *de P est*

$$\min\{k \text{ tel que } P \mid X^k - 1\}. \tag{4.103}$$

Définition 4.114 (Polynôme primitif)**.**
Soit p, un nombre premier et P un polynôme de degré n dans $\mathbb{F}_p[X]$. Nous disons que P est ***primitif*** *si*

(1) P est unitaire et irréductible,

(2) les racines de P sont d'ordre $p^n - 1$ dans $\mathbb{F}_p[X]/P$.

Définition 4.115 (Polynôme primitif au sens du pgcd)**.**
Soit un anneau A. Un polynôme $P \in A[X]$ est ***primitif au sens du pgcd*** *si ses coefficients sont premiers entre eux.*

4.116.
Pour rappel, il y a plusieurs façons de périphraser le fait que les coefficients soient premiers entre eux. Nous pouvons dire ...

(1) Le pgcd de ses coefficients est 1 parce que c'est la définition 4.74 d'avoir des nombres premiers entre eux.

(2) Le contenu de ses coefficients est 1. Parce que le contenu est précisément le pgcd, définition 4.112.

La notion de polynôme primitif au sens du pgcd est particulière aux polynôme à coefficients dans un anneau comme le montre le lemme suivant.

Lemme 4.117.
Si $\mathbb{K}$ est un corps, tout polynôme unitaire dans $\mathbb{K}[X]$ non nul est primitif au sens du pgcd.

Démonstration. Un polynôme unitaire a un 1 parmi ses coefficients, donc le pgcd est forcément 1. □

Lorsque nous utiliserons la notion de polynôme primitif au sens du pgcd, nous le mentionnerons explicitement. C'est pas exemple le cas pour le corollaire 4.124.

4.10.3 Racines des polynômes

Définition 4.118.
Soit A un anneau, et $P \in A[X]$. On appelle ***racine*** *un élément $\alpha \in A$ tel que $P(\alpha) = 0$; c'est-*

à-dire que, en remplaçant toutes les occurrences de X par α dans l'expression de P, on obtient 0.

Proposition 4.119.
Soit A un anneau et P un polynôme non nul dans $A[X]$. Si $\alpha \in A$ est une racine de P alors $X-\alpha$ divise P, et réciproquement.

Démonstration. Nous notons le polynôme $\mu = X - \alpha$ par analogie avec le polynôme minimal dont il sera question dans la très semblable proposition 5.65. Le sens réciproque est clair : si μ divise P, alors α est racine de P.

Pour le sens direct, remarquons que si α est racine de P, alors P est de degré au moins égal à 1, et nous pouvons donc effectuer la division euclidienne [25] de P par μ : il existe des polynômes Q et R tels que

$$P = Q\mu + R \tag{4.104}$$

avec $\deg(R) < \deg(\mu)$. Donc R est une constante, élément de A : appelons-le a. En évaluant (4.104) en α, il vient

$$0 = P(\alpha) = Q(\alpha)\mu(\alpha) + a, \tag{4.105}$$

et nous en déduisons que $a = 0$, ce qui montre que $P = Q\mu$ et que μ divise P. □

Définition 4.120 (Racine simple et multiple d'un polynôme)**.**
*Soit A un anneau ainsi qu'un polynôme $P \in A[X]$ et $\alpha \in A$ racine de P. La **multiplicité** de α par rapport à P est l'entier h tel que P est divisible par $(X-\alpha)^h$ mais pas divisible par $(X-\alpha)^{h+1}$. Nous noterons $\theta_\alpha(P)$ la multiplicité de α par rapport à P.*

Pour une définition générale d'une racine simple de l'équation $f(x) = 0$, voir la définition 32.49.

La proposition 4.119 nous indique que toute racine est de multiplicité au moins 1.

Proposition 4.121.
L'élément $\alpha \in A$ est une racine de multiplicité h du polynôme P si et seulement s'il existe $Q \in A[X]$ tel que $P = (X-\alpha)^h Q$ avec $Q(\alpha) \neq 0$.

Lemme 4.122.
Soient P et Q des polynômes non nuls de $A[X]$ et $\alpha \in A$. Alors

(1) $\theta_\alpha(P+Q) \leqslant \min\{\theta_\alpha(P), \theta_\alpha(Q)\}$, et l'égalité a lieu si $\theta_\alpha(P) \neq \theta_\alpha(Q)$;

(2) $\theta_\alpha(PQ) \geqslant \theta_{\alpha(P)} + \theta_\alpha(Q)$, et l'égalité a lieu si A est intègre.

Théorème 4.123.
Soit A un anneau intègre et $P \in A[X]\backslash\{0\}$, un polynôme de degré n. Si $\alpha_1, \ldots, \alpha_p \in A$ sont des racines deux à deux distinctes de multiplicités $k_1, \ldots, k_p$, alors il existe $Q \in A[X]$, de degré $n-p$, tel que $P = Q\prod_{i=1}^{p}(X-\alpha_i)^{k_i}$ et $Q(\alpha_i) \neq 0$ pour tout i. De plus la somme des multiplicités des racines de P est au plus $\deg(P)$.

Démonstration. Si $p = 1$, soit α une racine de multiplicité k de P. La définition de la multiplicité d'une racine nous dit que P est divisible par $(X-\alpha)^k$ mais pas par $(X-\alpha)^{k+1}$. Donc il existe $Q \in \mathbb{A}[X]$ tel que $P = Q(X-\alpha)^k$. Il reste à voir que $Q(\alpha) \neq 0$. Cela est une conséquence de la proposition 4.119 : si $Q(\alpha)$ était nul, on pourrait lui factoriser $(X-\alpha)$ et donc avoir $(X-\alpha)^{k+1}$ qui se factorise dans P, ce qui n'est pas possible.

Nous supposons que $p \geqslant 2$ et nous effectuons une récurrence sur p. Nous considérons donc les $p-1$ premières racines $\alpha_1, \ldots, \alpha_{p-1}$ et un polynôme $R \in \mathbb{A}[X]$ tel que $R(\alpha_i) \neq 0$ pour $i = 1, \ldots, p-1$ et

$$P = \underbrace{(X-\alpha_1)^{k_1} \ldots (X-\alpha_{p-1})^{k_{p-1}}}_{S} R. \tag{4.106}$$

25. Théorème 4.106.

Par hypothèse $P(\alpha_p) = S(\alpha_p)R(\alpha_p) = 0$. L'anneau $\mathbb{A}$ étant intègre, $S(\alpha_p) \neq 0$ parce que $\alpha_i \neq \alpha_p$ pour $i \neq p$. Par conséquent, $R(\alpha_p) = 0$.

Nous devons encore vérifier que la multiplicité α_p est k_p par rapport à R. Pour cela nous utilisons le point (2) du lemme 4.122 afin de dire que le degré de α_p pour $P = SR$ est k_p. Par conséquent

$$R = (X - \alpha_p)^{k_p} T \tag{4.107}$$

avec $T(\alpha_p) \neq 0$ et enfin

$$P = \prod_{i=1}^{p} (X - \alpha_i) T. \tag{4.108}$$

De plus $T(\alpha_i) \neq 0$, sinon $R(\alpha_i)$ serait nul. □

Corollaire 4.124 (Conséquence du lemme de Gauss[37])**.**
Soit A un anneau factoriel et Frac(A) *son corps des fractions. Un polynôme non constant $P \in A[X]$ est irréductible (sur A) si et seulement s'il est irréductible et primitif au sens du pgcd*[26] *sur* Frac$(A)[X]$.

4.10.4 Quelques identités

Lemme 4.125.
Quelques identités de polynômes.

(1) Si n est impair, alors $1 + X$ divise $1 + X^n$.

(2) Pour tout n nous avons $X^n - 1 = (X - 1)(1 + X + \cdots + X^{n-1})$.

(3) $X^n - a^n = (X - a) \sum_{i=0}^{n-1} a^i X^{n-1-i}$.

Démonstration. Nous démontrons uniquement le point (2), puisque les autres ont été vus en début de chapitre[27]. Le cas $n = 1$ est évident. Procédons alors par récurrence en considérant un nombre entier impair n :

$$\begin{aligned} 1 + X^{n+2} &= 1 + X^n + X^{n+2} - X^n && (4.109a)\\ &= (1 + X)P + X^n(X^2 - 1) && (4.109b)\\ &= (1 + X)P + X^n(X + 1)(X - 1) && (4.109c)\\ &= (1 + X)\big(P + X^n(X - 1)\big). && (4.109d) \end{aligned}$$

□

26. Définition 4.115.
27. Voir l'égalité (4.5).

Chapitre 5

Corps

5.1 Généralités

5.1.
Nous trouvons parfois le terme **anneau à division**. Cela provient du fait que dans beaucoup de cas on considère uniquement des corps commutatifs ; donc on voudrait une façon de parler d'un anneau dont tous les éléments non nuls sont inversibles. Dans ce cadre on dit :

— Un anneau à division est un anneau dont tous les éléments non nuls sont inversibles,
— Un corps est un anneau à division commutatif.

Pour prendre un exemple de cette différence, le théorème de Wedderburn 5.128 est énoncé ici sous les termes «Tout corps fini est commutatif». Sous-entendu : la commutativité ne fait pas partie de la définition d'un corps. Par contre dans [2] il est énoncé sous les termes «Tout anneau à division fini est un corps». Chez lui, un corps est toujours commutatif et un anneau à division est ce que nous appelons ici un corps.

5.1.1 Corps ordonnés

Nous avons vu la définition de corps totalement ordonné en 2.25.

Définition 5.2 ([38])**.**
Un corps est ***formellement réel*** *si* -1 *n'est pas une somme de carrés.*

Proposition 5.3.
Un corps totalement ordonné est formellement réel.

Démonstration. Soit un corps totalement ordonné $(\mathbb{K}, \leqslant)$ et $a \in \mathbb{K}$ alors $a^2 \geqslant 0$. En effet si $a \geqslant 0$ alors $a^2 = a \times a \geqslant 0$ directement par la définition 2.25(1)b. Si $a \leqslant 0$ alors $-a \geqslant 0$ et

$$a^2 = (-a)^2 \geqslant 0. \tag{5.1}$$

Vu que $-1 < 0$, il ne peut donc pas être écrit comme un carré. A fortiori comme somme de carrés. □

5.1.2 Automorphismes de $\mathbb{R}$ et $\mathbb{C}$

Proposition 5.4 ([39, 1])**.**
L'identité est l'unique automorphisme du corps $\mathbb{R}$*.*

Démonstration. Soit un automorphisme $\sigma : \mathbb{R} \to \mathbb{R}$. Comme pour tout automorphisme,

$$\sigma(a) = \sigma(1a) = \sigma(1)\sigma(a). \tag{5.2}$$

Donc $\sigma(1) = 1$.

Identité sur les rationnels De plus

$$\sigma(n) = \sigma(1 + \ldots + 1) = \sigma(1) + \ldots + \sigma(1) = n, \tag{5.3}$$

et

$$\sigma\left(\frac{1}{n}\right) + \ldots + \sigma\left(\frac{1}{n}\right) = \sigma\left(\frac{1}{n} + \ldots + \frac{1}{n}\right) = \sigma(1) = 1. \tag{5.4}$$

Donc $\sigma(1/n) = 1/n$.

Nous en déduisons que pour tout $q \in \mathbb{Q}$, $\sigma(q) = q$. Cela ne suffit pas pour déduire $\sigma(x) = x$ pour tout $x \in \mathbb{R}$ parce que rien n'indique que σ soit continue.

Positive sur les positifs Si $x > 0$ alors $\sigma(x) = \sigma(\sqrt{x})^2 > 0$.

Croissance Si $x > y$ alors $x - y > 0$ et $\sigma(x - y) > 0$. Cela donne $\sigma(x) > \sigma(y)$.

Identité sur les réels Soit un irrationnel $x \in \mathbb{R}$ et une suite $q_i \in \mathbb{Q}$ qui converge de façon croissante vers x. Soit $\epsilon > 0$ dans $\mathbb{Q}$. Il existe N tel que si $i > N$ alors $q_i + \epsilon > x$; en appliquant σ à cette inégalité et en se souvenant que σ est l'identité sur $\mathbb{Q}$,

$$q_i + \epsilon > \sigma(x). \tag{5.5}$$

Mais de plus, $q_i < x$ donne $\sigma(q_i) < \sigma(x)$, c'est à dire $q_i < \sigma(x)$. En regroupant ces deux inégalités,

$$q_i < \sigma(x) < q_i + \epsilon \tag{5.6}$$

pour tout $\epsilon > 0$ dans $\mathbb{Q}$ et $i > N$. Ce ϵ étant fixé nous pouvons prendre la limite des inégalités (5.6) :

$$x \leqslant \sigma(x) \leqslant x + \epsilon. \tag{5.7}$$

Cela étant valable pour tout $\epsilon > 0$ dans $\mathbb{Q}$, nous avons bien $x = \sigma(x)$.

□

Remarque 5.5.
Certains[39] pensent que l'énoncé de cette proposition, ne parlant que de *corps* $\mathbb{R}$ n'autorise pas l'utilisation d'autre structure réelle que celle de corps. Du coup il faut reconstruire la notion d'ordre à partir seulement du langage des corps. Par exemple en disant que $a > b$ si et seulement si il existe k tel que $a = b + k^2$.

On peut s'en sortir en donnant l'énoncé suivant : «Si $\mathbb{K}$ est un corps isomorphe (en tant que corps) à $\mathbb{R}$ alors son unique automorphisme est l'identité». Cela se démontre immédiatement en disant que si f est un automorphisme de $\mathbb{K}$ et si ϕ est un isomorphisme $\mathbb{K} \to \mathbb{R}$ alors $\phi \circ f \circ \phi^{-1}$ est un automorphisme de $\mathbb{R}$. Donc il est l'identité et f l'est également.

Attention cependant à prouver que ϕ^{-1} est un morphisme. En effet en posant $\phi^{-1}(x) = a$ et $\phi^{-1}(y) = b$ nous avons

$$\phi\big(\phi^{-1}(x) + \phi^{-1}(y)\big) = x + y \tag{5.8}$$

parce que ϕ est un morphisme. D'autre part,

$$\phi\big(\phi^{-1}(x) + \phi^{-1}(y)\big) = \phi(a + b). \tag{5.9}$$

Donc

$$\phi^{-1}(x + y) = \phi^{-1}\big(\phi(a) + \phi(b)\big) = \phi^{-1}\big(\phi(a + b)\big) = a + b = \phi^{-1}(x) + \phi^{-1}(y). \tag{5.10}$$

Proposition 5.6.
Un automorphisme du corps $\mathbb{C}$ qui fixe $\mathbb{R}$ est soit l'identité soit la conjugaison complexe[1].

Démonstration. Soit un automorphisme σ vérifiant la condition de fixer $\mathbb{R}$. Alors la restriction de σ à $\mathbb{R}$ est un automorphisme de $\mathbb{R}$ et y est donc l'identité par la proposition 5.4.

En ce qui concerne les imaginaires purs,

$$-1 = \sigma(-1) = \sigma(ii) = \sigma(i)^2. \tag{5.11}$$

Donc $\sigma(i)$ est un élément de $\mathbb{C}$ vérifiant $\sigma(i)^2 = -1$. C'est à dire $\sigma(i) = \pm i$.

Si $\sigma(i) = i$ alors $\sigma = \mathrm{Id}$. Si $\sigma(i) = -i$ alors σ est la conjugaison complexe. □

1. Par «fixer $\mathbb{R}$» nous entendons que $\sigma(\mathbb{R}) = \mathbb{R}$, pas spécialement que $\sigma(x) = x$ pour tout $x \in \mathbb{R}$

5.1.3 Corps premier

Définition 5.7.
Un corps est ***premier*** *s'il est son seul sous corps. Le* ***sous corps premier*** *d'un corps est l'intersection de tous ses sous corps.*

Lemme 5.8.
Un corps premier est commutatif

Démonstration. Le centre d'un corps est certainement un sous corps. Par conséquent un corps premier doit être contenu dans son propre centre, c'est à dire être commutatif. □

Définition 5.9.
Soit p un nombre premier. Nous notons $\mathbb{F}_p = \mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$.

Nous verrons plus loin (section 5.7) comment nous pouvons définir $\mathbb{F}_{p^n}$ lorsque p est premier, ainsi que l'unicité d'un tel corps.

Nous avons par exemple

$$\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z} = \{0, 1\} \tag{5.12}$$

avec la loi $2 = 0$.

Notons que $\mathbb{F}_p$ est un corps possédant p éléments. L'ensemble $\mathbb{F}_p^*$ est un groupe d'ordre $p - 1$.

Lemme 5.10.
Les corps $\mathbb{Q}$ et $\mathbb{F}_p$ (avec p premier) sont premiers.

Démonstration. Tout sous corps de $\mathbb{Q}$ doit contenir 1, et par conséquent $\mathbb{Z}$. Devant également contenir tous les inverses, il contient $\mathbb{Q}$.

Tout sous corps de $\mathbb{F}_p$ doit contenir 1 et donc $\mathbb{F}_p$ en entier. Par ailleurs nous savons de la proposition 4.54 que $\mathbb{F}_p$ est un corps lorsque p est premier. □

Proposition 5.11.
Soit $\mathbb{K}$ un corps de caractéristique p et $\mathbb{P}$ son sous corps premier. Si $p = 0$ alors $\mathbb{P} = \mathbb{Q}$. Si $p > 0$, alors $\mathbb{P} = \mathbb{F}_p$.

Démonstration. Notons d'abord que la caractéristique d'un corps est toujours un nombre premier parce qu'un corps est en particulier un anneau intègre (proposition 4.43).

Étant donné que 1 est dans tout sous corps, nous devons avoir $\mathbb{Z}1 \subseteq \mathbb{P}$. Si $p = 0$, alors $\mathbb{Z}1 \simeq \mathbb{Z}$, et nous avons

$$\mathbb{Z}1_{\mathbb{A}} \subset \mathbb{P} \subset \mathbb{K}. \tag{5.13}$$

Pour chaque $(n, m) \in \mathbb{Z}1_{\mathbb{A}} \times (\mathbb{Z}1_{\mathbb{A}})^*$ l'élément $nm^{-1} \in \mathbb{K}$ est dans $\mathbb{P}$ parce que $\mathbb{P}$ est un corps. Nous en déduisons que le corps des fractions de $\mathbb{Z}$ est contenu dans $\mathbb{P}$ par conséquent $\mathbb{P} = \mathbb{Q}$ (théorème 4.57).

Si par contre la caractéristique de $\mathbb{K}$ est $p \neq 0$, nous avons $\mathbb{Z}1_{\mathbb{A}} \simeq \mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ par le lemme 4.20. L'ensemble $\mathbb{F}_p$ étant un corps, c'est le corps premier de $\mathbb{K}$. □

Proposition 5.12.
Soit $\mathbb{K}$ un corps et $\mathbb{P}$ sont sous corps premier. Si $\sigma \in \mathrm{Aut}(\mathbb{K})$ alors $\sigma|_{\mathbb{P}} = \mathrm{Id}$, c'est à dire que $\sigma(x) = x$ pour tout $x \in \mathbb{P}$.

5.1.4 Petit théorème de Fermat

Théorème 5.13 (Petit théorème de Fermat)**.**
Soit p un nombre premier. Si $x \in \mathbb{F}_p$ alors $x^p = x$. Si $x \in (\mathbb{F}_p)^*$, alors $x^{p-1} = 1$.

En particulier si $x \in \mathbb{F}_p^*$ alors $x^{-1} = x^{p-2}$.

Démonstration. Étant donné que $\mathbb{F}_p$ est un corps commutatif et que p est premier, la proposition 4.23 nous indique que $\sigma(x) = x^p$ est un automorphisme. La proposition 5.12 nous indique alors que

$$a^p = a. \tag{5.14}$$

Si a est inversible alors $a^{p-1} = a^p a^{-1} = 1$. □

Remarque 5.14.
Une autre façon d'énoncer le petit théorème de Fermat 5.13 est que si p est premier et si a est premier avec a, alors $a^{p-1} \in [1]_p$. Le nombre a n'est pas premier avec p uniquement lorsque a est multiple de p. Dans ce cas c'est $a = 0$ dans $\mathbb{F}_p$ et donc $a^{p-1=0}$.

Exemple 5.15
Soit $\mathbb{K} = \mathbb{F}_{29}^*$. Le nombre 29 étant premier, $\mathbb{K}$ est un corps premier. C'est le corps des entier modulo 29. Nous avons donc

$$-142 = -113 = -84 = -55 = -26 = 3 = 32 = 61 = 90 = 119. \tag{5.15}$$

Le petit théorème de Fermat nous permet aussi de calculer des exposants et des inverses. Nous avons

$$x^{-a} = (x^a)^{27} = x^{27a}. \tag{5.16}$$

Le nombre $27a$ peut être grand par rapport à 29. Étant donné que $1 = x^{28}$ nous avons

$$x^{-a} = x^{27a \mod 28}. \tag{5.17}$$

Dans les exposants nous calculons modulo 28. △

5.2 Théorème des deux carrés

Proposition 5.16.
Soit p un nombre premier et P un élément de $\mathbb{F}_p[X]$. L'anneau $\mathbb{F}_p[X]/(P)$ est intègre si et seulement si P est irréductible dans $\mathbb{F}_p[X]$.

Démonstration. Supposons que P soit réductible dans $\mathbb{F}_p[X]$, c'est à dire qu'il existe $Q, R \in \mathbb{F}_p[X]$ tels que $P = QR$. Dans ce cas, $\bar{Q}$ est diviseur de zéro dans $\mathbb{F}_p[X]/(P)$ parce que $\bar{Q}\bar{R} = 0$.

Nous supposons maintenant que $\mathbb{F}_p[X]/(P)$ ne soit pas intègre : il existe des polynômes $R, Q \in \mathbb{F}_p[X]$ tels que $\bar{Q}\bar{R} = 0$. Dans ce cas le polynôme QR est le produit de P par un polynôme : $QR = PA$. Tous les facteurs irréductibles de A étant soit dans Q soit dans R, il est possible de modifier un peu Q et R pour obtenir $QR = P$, ce qui signifie que P n'est pas irréductible. □

5.2.1 Un peu de structure dans $\mathbb{Z}[i]$

Lemme 5.17.
L'application

$$\begin{aligned} N\colon \mathbb{Z}[i] &\to \mathbb{N} \\ a + bi &\mapsto a^2 + b^2 \end{aligned} \tag{5.18}$$

est un stathme euclidien pour $\mathbb{Z}[i]$.

Démonstration. Soient $t, t \in \mathbb{Z}[i]\backslash\{0\}$ et

$$\frac{z}{t} = x + iy \tag{5.19}$$

dans $\mathbb{C}$. Nous considérons $q = a + bi$ où a et b sont les entiers les plus proches de x et y. S'il y a *ex aequo*, on prend au hasard [2]. Alors nous avons

$$|\frac{z}{t} - q| \leqslant \frac{|1+i|}{2} = \frac{\sqrt{2}}{2} < 1. \tag{5.20}$$

On pose $r = z - qt$ qui est bien un élément de $\mathbb{Z}[i]$. De plus

$$|r| = |z - qt| = |t||\frac{z}{t} - q| < |t|, \tag{5.21}$$

c'est à dire que $|r|^2 < |t|^2$ et donc $N(r) < N(t)$. □

Étant donné que $\mathbb{Z}[i]$ est euclidien, il est principal (proposition 4.84).

Lemme 5.18.
Les éléments inversibles de $\mathbb{Z}[i]$ sont $\{\pm 1, \pm i\}$.

Démonstration. Déterminons les éléments inversibles de $\mathbb{Z}[i]$. Si $z \in \mathbb{Z}[i]^*$, alors il existe $z' \in \mathbb{Z}[i]^*$ tel que $zz' = 1$. Dans ce cas nous aurions

$$1 = N(zz') = N(z)N(z'), \tag{5.22}$$

ce qui est uniquement possible avec $N(z) = N(z') = 1$, c'est à dire $z = \pm 1$ ou $z = \pm i$. Nous avons donc

$$\mathbb{Z}[i]^* = \{\pm 1, \pm i\}. \tag{5.23}$$

□

Nous notons $\Sigma = \{a^2 + b^2 \text{ tel que } a, b \in \mathbb{N}\}$.

Lemme 5.19.
L'ensemble Σ est un sous-monoïde de $\mathbb{N}$.

Démonstration. Il suffit de prouver que si $m, n \in \Sigma$, alors le produit mn est également dans Σ. Si N est le stathme euclidien sur $\mathbb{Z}[i]$, alors $n \in \Sigma$ si et seulement s'il existe $z \in \mathbb{Z}[i]$ tel que $N(z) = n$. Si $z, z' \in \mathbb{Z}[i]$, alors $zz' \in \mathbb{Z}[i]$ et de plus

$$N(zz') = N(z)N(z') = nm. \tag{5.24}$$

Donc nm est l'image de zz' par N, ce qui prouve que $nm \in \Sigma$. □

Théorème 5.20 (Théorème des deux carrés, version faible)**.**
Un nombre premier est somme de deux carrés si et seulement si $p = 2$ ou $p \in [1]_4$.

Remarque 5.21.
Il n'est pas dit que les nombres dans $[1]_4$ sont premiers ($9 = 8 + 1$ ne l'est pas par exemple). Le théorème signifie que (à part 2), si un nombre premier est dans $[1]_4$ alors il est somme de deux carrés, et inversement, si un nombre premier est somme de deux carrés, il est dans $[1]_4$.

Démonstration. Soit p un nombre premier dans Σ. Si $a = 2k$, alors $a^2 = 4k^2$ et $a^2 = 0 \mod 4$. Si au contraire a est impair, $a = 2k + 1$ et $a^2 = 4k^2 + 1 + 4k = 1 \mod 4$. La même chose est valable pour b. Par conséquent, $a^2 + b^2$ est automatiquement $[0]_4$, $[1]_4$ ou $[2]_4$. Évidemment les nombres de la forme $0 \mod 4$ ne sont pas premiers ; parmi les $2 \mod 4$, seul $p = 2$ est premier (et vaut $1^2 + 1^2$).

Nous avons démontré que les seuls premiers de la forme $a^2 + b^2$ sont $p = 2$ et les $p = 1 \mod 4$. Il reste à faire le contraire : démontrer que si un nombre premier p vaut $1 \mod 4$, alors il est premier. Nous considérons l'anneau

$$\mathbb{Z}[i] = \{a + bi \text{ tel que } a, b \in \mathbb{Z}\}. \tag{5.25}$$

2. Dans l'exemple 4.83, nous prenions toujours l'inférieur parce que le stathme tenait compte de la positivité.

puis l'application

$$\begin{aligned} N\colon \mathbb{Z}[i] &\to \mathbb{N} \\ a+bi &\mapsto a^2+b^2. \end{aligned} \tag{5.26}$$

Un peu de calcul dans $\mathbb{C}$ montre que pour tout $z, z' \in \mathbb{Z}[i]$, $N(zz') = N(z)N(z')$.

Nous savons que les éléments inversibles de $\mathbb{Z}[i]$ sont ± 1 et $\pm i$ (lemme 5.18).

Le lemme 5.17 montre que $\mathbb{Z}[i]$ est un anneau euclidien parce que N est un stathme. L'anneau $\mathbb{Z}[i]$ étant euclidien, il est principal (proposition 4.84).

Pour la suite, nous allons d'abord montrer que $p \in \Sigma$ si et seulement si p n'est pas irréductible dans $\mathbb{Z}[i]$, puis nous allons voir quels sont les irréductibles de $\mathbb{Z}[i]$.

Soit p, un nombre premier dans Σ. Si $p = a^2 + b^2$, alors nous avons $p = (a + ib)(a - bi)$, mais étant donné que p est premier, nous avons $a \neq 0$ et $b \neq 0$. Du coup p n'est pas inversible dans $\mathbb{Z}[i]$, mais il peut être écrit comme le produit de deux non inversibles. Le nombre p est donc non irréductible dans $\mathbb{Z}[i]$.

Dans l'autre sens, nous supposons que p est un nombre premier non irréductible dans $\mathbb{Z}[i]$. Nous avons alors $p = zz'$ avec ni z ni z' dans $\{\pm 1, \pm i\}$. En appliquant N nous avons

$$p^2 = N(p) = N(z)N(z'). \tag{5.27}$$

Vu que p est premier, cela est uniquement possible avec $N(z) = N(z') = p$ (avoir $N(z) = 1$ est impossible parce que cela dirait que z est inversible). Si $z = a + ib$, alors $p = N(z) = a^2 + b^2$, et donc $p \in \Sigma$.

Nous savons déjà que $\mathbb{Z}[i]$ est un anneau principal et n'est pas un corps ; la proposition 4.70 s'applique donc et p sera non irréductible si et seulement si l'idéal (p) sera non premier. Le fait que (p) soit un idéal non premier implique que le quotient $\mathbb{Z}[i]/(p)$ est non intègre (c'est la définition d'un idéal premier). Nous cherchons donc les nombres premiers pour lesquels le quotient $\mathbb{Z}[i]/(p)$ n'est pas intègre.

Nous commençons par écrire le quotient $\mathbb{Z}[i]/(p)$ sous d'autres formes. D'abord en remarquant que si I et J sont deux idéaux, on a $(\mathbb{A}/I)/J \simeq (\mathbb{A}/J)/I$, du coup, en tenant compte du fait que $\mathbb{Z}[i] = \mathbb{Z}[X]/(X^2 + 1)$, nous avons

$$\mathbb{Z}[i]/(p) = (\mathbb{Z}[X]/(p))/(X^2 + 1) = \mathbb{F}_p[X]/(X^2 + 1). \tag{5.28}$$

Nous avons donc équivalence des propositions suivantes :

$$p \in \Sigma \tag{5.29a}$$

$$\mathbb{F}_p[X]/(X^2 + 1) \text{ n'est pas intègre} \tag{5.29b}$$

$$X^2 + 1 \text{ n'est pas irréductible dans } \mathbb{F}_p \tag{5.29c}$$

$$X^2 + 1 \text{ admet une racine dans } \mathbb{F}_p \tag{5.29d}$$

$$-1 \in (\mathbb{F}_p^*)^2 \tag{5.29e}$$

$$\exists y \in \mathbb{F}_p^* \text{ tel que } y^2 = -1. \tag{5.29f}$$

Le point (5.29c) vient de la proposition 5.16. Maintenant nous utilisons le fait que p soit un premier impair (parce que le cas de $p = 2$ est déjà complètement traité), donc $(p-1)/2 \in \mathbb{N}$ et nous avons, pour le y de la dernière ligne,

$$(-1)^{(p-1)/2} = (y^2)^{(p-1)/2} = y^{p-1} = 1 \tag{5.30}$$

parce que dans $\mathbb{F}_p$ nous avons $y^{(p-1)} = 1$ par le petit théorème de Fermat (théorème 5.13). Du coup p doit vérifier

$$1 = (-1)^{(p-1)/2}, \tag{5.31}$$

c'est à dire $\frac{p-1}{2} = 0 \mod 2$ ou encore $p = 1 \mod 4$. □

Théorème 5.22 (Théorème des deux carrés[2]).
Soit $n \geqslant 2$ un nombre dont nous notons

$$n = \prod_{p\in\mathcal{P}} p^{v_p(n)} \tag{5.32}$$

où $\mathcal{P}$ est l'ensemble des nombres premiers. Alors $n \in \Sigma$ si et seulement si pour tout $p \in \mathcal{P} \cap [3]_4$, nous avons $v_p(n) \in [0]_2$ (c'est à dire $v_p(n)$ est pair).

Démonstration. **Condition suffisante.** Le produit (5.32) est évidemment un produit fini que nous pouvons alors regrouper en quatre parties : $\mathcal{P} \cap [0]_4$, $\mathcal{P} \cap [1]_4$, $\mathcal{P} \cap [2]_4$ et $\mathcal{P} \cap [3]_4$.

— Il n'y a pas de nombres premiers dans $[0]_4$.
— Les nombres premiers de $[1]_4$ sont dans Σ. Le produit d'éléments de Σ étant dans Σ, nous avons

$$\prod_{p\in\mathcal{P}\cap[1]_4} p^{v_p(n)} \in \Sigma. \tag{5.33}$$

— Le seul nombre premier dans $[2]_4$ est 2. C'est un élément de Σ.
— Le produit

$$\prod_{p\in\mathcal{P}\cap[3]_4} p^{v_p(n)} \tag{5.34}$$

est par hypothèse un produit de carrés ($v_p(n)$ est pair), et est donc un carré.

Au final le produit $\prod_{p\in\mathcal{P}} p^{v_p(n)}$ est un produit d'un carré par un élément de Σ, ce qui est encore un élément de Σ.

Pour cette partie, nous avons utilisé et réutilisé le lemme 5.19.

Condition nécessaire. Soit p, un nombre premier. Nous voulons montrer que

$$\{v_p(n) \text{ tel que } n \in \Sigma\} \subset [2]_2. \tag{5.35}$$

Pour montrer cela nous allons procéder par récurrence sur les ensembles

$$E_k = \{v_p(n) \text{ tel que } n \in \Sigma\} \cap \{0, \dots, k\}. \tag{5.36}$$

Il est évident que les éléments de E_0 sont pairs, vu qu'il n'y a que zéro, qui est pair.

Supposons que $E_k \subset [0]_2$, et montrons que $E_{k+1} \subset [0]_2$. Soit un élément de E_{k+1}, c'est à dire $v_p(n) \leqslant k+1$ avec $n = a^2 + b^2$. Si $v_p(n) = 0$ alors l'affaire est réglée ; sinon c'est que p divise n. Mais dans $\mathbb{Z}[i]$ nous avons

$$n = a^2 + b^2 = (a + bi)(a - bi) \tag{5.37}$$

Vu que $\mathbb{Z}[i]$ est principal, le lemme de Gauss 4.78 nous dit que si p divise n, alors il doit diviser soit $a + bi$, soit $a - bi$ (et du coup en fait les deux). Nous avons alors $p \mid a$ et $p \mid b$ en même temps. Du coup

$$p^2 \mid a^2 + b^2 = n. \tag{5.38}$$

Posons $a = pa'$ et $b = pb'$ avec $a', b' \in \mathbb{N}$. Nous avons

$$\frac{n}{p^2} = \frac{p^2a'^2 + p^2b'^2}{p^2} = a'^2 + b'^2 \in \Sigma. \tag{5.39}$$

Mais par construction,

$$v_p\left(\frac{n}{p^2}\right) = v_p(n) - 2 < k. \tag{5.40}$$

Donc $v_p(\frac{n}{p^2})$ est pair et du coup $v_p(n)$ doit également être pair.

□

5.2.2 Résultats chinois

Nous allons maintenant parler du système d'équations

$$\begin{cases} x = a_1 \mod p & (5.41a) \\ x = a_2 \mod q & (5.41b) \end{cases}$$

avec a_1, a_2 donnés dans $\mathbb{Z}$ et p, q des entiers premiers entre eux. Le lemme chinois nous donne la liste des solutions ainsi qu'une manière de les construire. Le théorème chinois en sera une espèce de corollaire qui établira l'isomorphisme d'anneaux $\mathbb{Z}/pq\mathbb{Z} \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$. Voir les-mathematiques.net.

Lemme 5.23 (Lemme chinois [40])**.**
Soient n_1, n_2 deux entiers premiers entre eux. Soient $a_1, a_2 \in \mathbb{Z}$. Les solutions du système

$$\begin{cases} x = a_1 \mod n_1 & (5.42a) \\ x = a_2 \mod n_2 & (5.42b) \end{cases}$$

pour $x \in \mathbb{Z}/n_1n_2\mathbb{Z}$ sont données de la façon suivante. Soient u_1, u_2 deux entiers qui satisfont la relation de Bézout[3]

$$u_1n_1 + u_2n_2 = 1, \tag{5.43}$$

et

$$a = \big(a_1u_2n_2 + a_2u_1n_1\big) \mod (n_1). \tag{5.44}$$

Alors $x = a \mod (n_1n_2)$.

Démonstration. Vérifions que le x donné par $x = a \mod (n_1n_2)$ est bien une solution. D'abord

$$\begin{aligned} a \mod n_2 &= a_1u_2n_2 \mod n_1 & (5.45a) \\ &= a_1(1 - u_1n_1) \mod n_1 & (5.45b) \\ &= a_1 \mod n_1 & (5.45c) \end{aligned}$$

où nous avons utilisé l'identité de Bézout (5.43). La vérification de $a \mod n_2 = a_2 \mod n_2$ est la même.

Soit maintenant $x \in \mathbb{Z}/n_1n_2\mathbb{Z}$ une solution du système (5.42) et a donné par la formule (5.44). Alors

$$\begin{aligned} (x - a) \mod n_1 &= \Big(a_1 - (a_1n_2u_2 + a_2u_1n_1)\Big) \mod n_1 & (5.46a) \\ &= a_1 - a_1u_2n_2 \mod n_1 & (5.46b) \\ &= 0, & (5.46c) \end{aligned}$$

donc $(x - a) \mod n_1 = 0$, ce qui signifie que $x - a$ est divisible par n_1. De la même façon, $(x - a) \mod n_2 = 0$ et $x - a$ est divisible par n_2. Nous savons maintenant que $x - a$ est divisible par n_1 et n_2. Étant donné que n_1 et n_2 sont premiers entre eux, nous en déduisons que $x - a$ est divisible par n_1n_2, ou encore que $x = a \mod n_1n_2$. □

Théorème 5.24 (Théorème chinois)**.**
Soient p, q deux naturels premiers entre eux. Si $p, q \geqslant 2$ alors l'application

$$\begin{aligned} \phi\colon \mathbb{Z}/pq\mathbb{Z} &\to \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z} \\ [x]_{pq} &\mapsto \big([x]_p, [x]_q\big) \end{aligned} \tag{5.47}$$

est un isomorphisme d'anneaux.

3. voir le théorème 3.32

Démonstration. Nous devons prouver que l'application ϕ respecte la somme, le produit et qu'elle est bijective. En ce qui concerne la somme,

$$\begin{aligned}\phi([q]_{pq}+[y]_{pq}) &= \phi\big((x+y) \mod pq\big) && (5.48a)\\ &= \big([x+y]_p,[x+y]_q\big) && (5.48b)\\ &= \big([x]_p+[y]_p,[x]_q+[y]_q\big) && (5.48c)\\ &= \big([x]_p,[x]_q\big)+\big([y]_p,[y]_q\big) && (5.48d)\\ &= \phi(x)+\phi(y). && (5.48e)\end{aligned}$$

En ce qui concerne le produit, c'est le même jeu : nous obtenons

$$\phi\big([xy]_{pq}\big) = \phi([x]_{pq}])\phi([y]_{pq}) \tag{5.49}$$

en utilisant le fait que $[xy]_p = [x]_p[y]_p$.

Montrons maintenant que ϕ est surjective. Soient $y_1, y_2 \in \mathbb{Z}$ et $x \in \mathbb{Z}$. Demander

$$\phi([x]_{pq}) = \big([y_1]_p,[y_2]_q\big) \tag{5.50}$$

revient à demander que $[x]_p = [y_1]_p$ et $[x]_q = [y_2]_q$, c'est à dire que x résolve le système

$$\begin{cases} x = y_1 \mod p & (5.51a)\\ x = y_2 \mod q. & (5.51b)\end{cases}$$

Le lemme chinois 5.23 nous assure qu'une solution existe.

En ce qui concerne l'injectivité, nous supposons que x et y soient deux entiers tels que

$$\phi([x]_{pq}) = \phi([y]_{pq}). \tag{5.52}$$

Nous en déduisons le système

$$\begin{cases} x \mod p = y \mod p & (5.53a)\\ x \mod q = y \mod q & (5.53b)\end{cases}$$

c'est à dire qu'il existe des entiers k et l tels que $x = y + kp$ et $x = y + lq$ ou encore tels que

$$kp + lq = 0. \tag{5.54}$$

Étant donné que p et q sont premiers entre eux, la seule possibilité est $k = l = 0$, c'est à dire $x = y$. □

Théorème 5.25 (Théorème chinois)**.**
Soit A un anneau commutatif, $n \geqslant 2$, des éléments $x_1, \dots, x_n$ dans A et des idéaux $I_1, \dots, I_n$ tels que $I_i + I_j = A$ pour tout $i \neq j$.

Alors il existe un $x \in A$ tel que $x - x_i \in I_i$ pour tout $1 \leqslant i \leqslant n$.

Démonstration. Pour $i \in \{1, \dots, n\}$ nous notons J_i le produit $J_i = \prod_{k\neq i} I_k$. Étant donné que chaque I_i est un idéal, nous avons $I_k \in J_i$ lorsque $i \neq k$.

Soit i fixé, et considérons $j \neq i$. Nous pouvons trouver $a_j \in I_i$ et $b_j \in I_j$ tel que $a_j + b_j = 1$. Nous avons alors

$$1 = \prod_{j\neq i}(a_j + b_j). \tag{5.55}$$

Par ailleurs $I_i + J_i = A$ parce que J_i contient I_k avec $k \neq i$ et $I_i + I_k = A$. Nous pouvons donc prendre $\alpha_i \in I_i$ et $\beta_i \in J_i$ tels que

$$\prod_{j\neq i}(a_j + b_j) = \alpha_i + \beta_i. \tag{5.56}$$

Nous considérons alors l'élément $x = \beta_1 x_1 + \cdots + \beta_n x_n$ et nous avons

$$x - x_1 = (\beta_1 - 1)x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{5.57a}$$
$$= -\alpha_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{5.57b}$$

Mais $\alpha_1 \in I_1$ et tous les autres termes sont dans les J_i avec $i \neq 1$. Par conséquent le tout est dans I_1. Ici nous utilisons par exemple le fait que $\beta_2 \in J_2 \subset I_1$ parce que les éléments de J_2 sont des produits d'éléments dont un facteur est dans I_1. □

Remarque 5.26.
Ce théorème chinois est bien une généralisation du lemme chinois 5.23. En effet, l'élément x dont il est question est solution du problème $x = x_i \mod I_i$. L'hypothèse $I_i + I_j = A$ n'est pas nouvelle non plus étant donné que si p et q sont des entiers premiers entre eux nous avons $p\mathbb{Z} + q\mathbb{Z} = \mathbb{Z}$ par le corollaire 3.33.

5.3 Polynômes à coefficients dans un corps

Nous supposons que $\mathbb{K}$ est une corps commutatif, et nous étudions l'anneau $\mathbb{K}[X]$.

Proposition 5.27.
L'anneau $\mathbb{K}[X]$ des polynômes sur un corps commutatif $\mathbb{K}$ est factoriel.

Le théorème suivant est une particularisation à $\mathbb{K}[X]$ du théorème chinois 4.75.

Théorème 5.28 (Théorème chinois)**.**
Si P et Q sont deux polynômes premiers entre eux, alors nous avons l'isomorphisme

$$\mathbb{K}[X]/(P, Q) \simeq \mathbb{K}[X]/(P) \times \mathbb{K}[X]/(Q). \tag{5.58}$$

5.3.1 Irréductibilité

Définition 5.29.
*Un polynôme est **irréductible** lorsqu'il ne peut pas être écrit sous la forme de produits de polynômes de degrés supérieurs à* 1.

Un polynôme est irréductible dans $\mathbb{K}[X]$ au sens de la définition 4.60 si et seulement s'il est irréductible au sens de la définition 5.29 parce que seules les constantes (non nulles) sont inversibles dans $\mathbb{K}[X]$.

Théorème 5.30 (d'Alembert-Gauss)**.**
Tout polynôme non constant à coefficients complexes possède au moins une racine complexe.

Exemple 5.31
Si un polynôme $P \in \mathbb{Z}[X]$ n'a que des racines complexes, ça ne l'empêche pas d'être réductible sur $\mathbb{Z}$. La réductibilité ne signifie pas qu'on peut mettre des racines en évidence. Par exemple le polynôme $P = X^4 + 3X^2 + 2$ est réductible sur $\mathbb{Z}$ en

$$P = (X^2 + 1)(X^2 + 2), \tag{5.59}$$

mais n'a pas de racines dans $\mathbb{Z}$. Par contre, il est vrai que si on veut réduire plus, il faut sortir de $\mathbb{Z}$.

△

Proposition 5.32.
Un polynôme irréductible à coefficients réels est soit de degré un soit de degré 2 *avec un discriminant négatif.*

Démonstration. Soit un polynôme P à coefficients réels de degré plus grand que 1. Alors le théorème de d'Alembert-Gauss (théorème 5.30) implique l'existence d'une racine $\alpha \in \mathbb{C}$. Si α est en réalité un réel, P est réductible. S'il ne l'est pas, alors il est facile de montrer que le conjugué complexe $\bar{\alpha}$ est également racine. Par conséquent les polynômes $(X-\alpha)$ et $(X-\bar{\alpha})$ divisent P dans $\mathbb{C}[X]$..

Ces deux polynômes sont premiers entre eux parce que

$$a(X-\alpha)+b(X-\bar{\alpha})=0 \tag{5.60}$$

implique $a=b=0$. Par conséquent le produit

$$X^2-(\alpha+\bar{\alpha})X+\alpha\bar{\alpha} \tag{5.61}$$

divise également P. Ce dernier est un polynôme à coefficients réels de degré 2. Donc tout polynôme de degré 3 ou plus est réductible. □

Définition 5.33.
*Nous disons que $P \in \mathbb{K}[X]$ est **scindé** sur $\mathbb{K}$ s'il est produit dans $\mathbb{K}[X]$ de polynômes de degré 1.*

Note : les constantes ne sont donc pas des polynôme scindés.

Proposition 5.34 (Critère d'Eisenstein)**.**
Soit le polynôme $P(X)=\sum_{k=0}^{n} a_n X^n$ dans $\mathbb{Z}[X]$. Nous supposons avoir un nombre premier p tel que

(1) p divise tous les $a_0,\dots,a_{n-1}$,

(2) p ne divise pas a_n,

(3) p^2 ne divise pas a_0.

Alors P est irréductible dans $\mathbb{Q}[X]$.

Si de plus P est primitif au sens du pgcd *(définition 4.115) alors P est irréductible dans $\mathbb{Z}[X]$.*

Démonstration. Nous considérons $\bar{P}$ le polynôme réduit modulo p, c'est à dire $\bar{P} \in \mathbb{F}_p[X]$. Étant donné que par hypothèse tous les coefficients sont multiples de p sauf a_n, nous avons $\bar{P}=cX^n$. Supposons par l'absurde que $P=QR$ avec $Q,R \in \mathbb{Q}[X]$. Alors le lemme de Gauss (4.78) impose $P,Q \in \mathbb{Z}[X]$.

Nous avons aussi, au niveau des réductions modulo p que $\bar{Q}\bar{R}=\bar{P}$. Or $\bar{P}$ est un monôme, donc $\bar{Q}$ et $\bar{R}$ doivent également l'être. Donc $\bar{Q}=dX^k$ et $\bar{R}=eX^{n-k}$ et en particulier $\bar{Q}(0)=\bar{R}(0)=0$, c'est à dire que $Q(0)$ et $R(0)$ sont divisibles par p. Cela impliquerait que $a_0=Q(0)R(0)$ soit divisible par p^2, ce qui est exclu par les hypothèses. Donc P est irréductible.

Supposons de surcroît que P est primitif au sens du pgcd. Il est donc irréductible et primitif sur $\mathbb{Q}[X]$ et le corollaire 4.124 nous dit alors que P est irréductible sur $\mathbb{Z}[X]$. □

Exemple 5.35
Soit le polynôme $P(X)=3X^4+15X^2+10$. Pour faire fonctionner le critère d'Eisenstein il nous faut un nombre premier p divisant 15 et 10, mais pas 3 et dont le carré ne divise pas 10. C'est vite vu que $p=5$ fait l'affaire. Le polynôme P est donc irréductible sur $\mathbb{Q}[X]$. △

5.3.2 Idéaux

Soit $P \in \mathbb{K}[X]$ un polynôme. Nous notons (P) l'idéal engendré par P :

$$(P)=\{PR \text{ tel que } R \in \mathbb{K}[X]\}. \tag{5.62}$$

Lemme 5.36.
Nous avons

(1) $(P) \subset (Q)$ si et seulement si Q divise P,

(2) $(P) = (Q)$ si et seulement si P et Q sont multiples (non nuls) l'un de l'autre.

Démonstration. Si $(P) \subset (Q)$, en particulier $P \in (Q)$ et il existe $R \in \mathbb{K}[X]$ tel que $P = QR$, ce qui signifie que Q divise P.

Si les idéaux de P et de Q sont identiques, l'un divise l'autre et l'autre divise l'un. Ils sont donc multiples l'un de l'autre. □

Théorème 5.37.
Soit $\mathbb{K}$ un corps commutatif.

(1) L'anneau $\mathbb{K}[X]$ est euclidien et principal.

(2) Si I est un idéal dans $\mathbb{K}[X]$ et si $P \in I$ est de degré minimal, alors $(P) = I$.

(3) De plus si $I \neq \{0\}$, il existe un unique polynôme unitaire μ tel que $I = (\mu)$.

Démonstration. Le point (1) a déjà été démontré dans le lemme 4.109 via le fait que $\mathbb{K}[X]$ est euclidien. Nous allons cependant donner ici une preuve directe que tous les idéaux de $\mathbb{K}[X]$ sont principaux. Si $I = \{0\}$, le résultat est évident. Nous supposons donc I non nul. Soit P de degré minimum parmi les éléments de I. Évidemment $(P) \subset I$. Nous allons démontrer qu'en réalité $(P) = I$.

Soit $P' \in I$. Par le théorème 4.106 de la division euclidienne[4], il existe Q et R dans $\mathbb{K}[X]$ tels que $P' = PQ + R$ avec $\deg(R) < \deg(P)$. Étant donné que $R = P' - PQ$ nous avons $R \in I$ et par conséquent $R = 0$ parce que P a été choisi de degré minimum dans I. Nous avons donc $P' = PQ$ et $I \subset (P)$.

L'existence d'un polynôme unitaire qui génère I est obtenu en choisissant $\mu = P/a_n$ où a_n est le coefficient du terme de plus haut degré. L'unicité d'un tel polynôme est obtenu par le fait que si μ et μ' génèrent le même idéal, alors ils sont multiples l'un de l'autre, or puisqu'ils sont unitaires, ils sont égaux. □

Nous voyons que n'importe quel polynôme de degré minimum dans un idéal génère l'idéal. Une importante conséquence du théorème 5.37 que nous verrons plus bas est que tout polynôme annulateur d'un endomorphisme est divisé par le polynôme minimal (proposition 9.110).

Corollaire 5.38.
Si $\mathbb{K}$ est un corps et si P est un polynôme irréductible, alors l'ensemble $\mathbb{L} = \mathbb{K}[X]/(P)$ est un corps. De plus $\mathbb{L}$ est un espace vectoriel de dimension $\deg(P)$.

Démonstration. En effet $\mathbb{K}[X]$ est un anneau principal par le théorème 5.37, par conséquent la proposition 4.71 déduit que $\mathbb{K}[X]/(P)$ est un corps.

Une base de $\mathbb{L}$ est donnée par les projections de $1, X, X^2, \ldots, X^{n-1}$. En effet ces élément forment une famille libre parce que si $\sum_{k=0}^{n-1} a_k \bar{X}^n = 0$ alors un représentant de cette classe doit être de la forme SP dans $\mathbb{K}[X]$, c'est à dire

$$\sum_{k=0}^{n-1} a_k X^k = SP, \tag{5.63}$$

ce qui n'est possible que si $S = 0$ et $a_k = 0$. D'autre part c'est un système générateur. En effet si $P = X^n + Q$ avec $\deg(Q) = n - 1$ alors

$$\bar{X}^{n+l} = \bar{X}^n \bar{X}^l = (\bar{P} - \bar{Q})\bar{X}^l = \bar{Q}\bar{X}^l. \tag{5.64}$$

Nous avons donc exprimé $\bar{X}^{n+l}$ comme une somme de termes de degré $n + l - 1$. Par récurrence nous pouvons exprimer tout $\bar{X}^{n+l}$ comme combinaison d'éléments de degré plus petit que n. □

Remarque 5.39.
Ce corollaire prendra une nouvelle jeunesse lorsque nous parlerons de polynômes d'endomorphismes, en particulier la proposition 9.120 va donner des précisions.

4. Ici $\mathbb{K}$ est un corps et donc l'hypothèse d'inversibilité est automatiquement vérifiée.

5.3.3 Bézout

Théorème 5.40 (Bézout)**.**
Les polynômes $P_1, \dots, P_n$ dans $\mathbb{K}[X]$ sont étrangers entre eux si et seulement s'il existe des polynômes $Q_1, \dots, Q_n \in \mathbb{K}[X]$ tels que

$$P_1Q_1 + \cdots + P_nQ_n = 1. \tag{5.65}$$

Deux polynômes P et Q ne sont donc pas premiers entre eux s'il existe des polynômes x et y tels que l'identité de Bézout soit vérifiée :

$$xP + yQ = 0; \tag{5.66}$$

cette dernière pourra être écrite en termes de la matrice de Sylvester, voir sous-section 9.1.8.

Lemme 5.41.
Soient $(P_i)_{i=1,\dots,n} \in \mathbb{K}[X]$ des polynômes étrangers deux à deux. Alors les polynômes

$$Q_i = \prod_{j \neq i} P_j \tag{5.67}$$

sont étrangers entre eux[5]*.*

Lemme 5.42 ([41])**.**
Soit $\mathbb{K}$ un corps commutatif et $A \subset \mathbb{K}$ un sous anneau de $\mathbb{K}$. Alors $A[X]$, vu comme idéal de $\mathbb{K}[X]$, est un idéal premier.

En d'autres termes, si $\phi \in \mathbb{K}[X]$, et s'il existe $Q \in \mathbb{A}[X]$ unitaire tel que $\phi Q \in \mathbb{A}[X]$, alors $\phi \in \mathbb{A}[X]$.

Lemme 5.43.
Quelques propriétés du PGCD dans les polynômes.

(1) Soit $P, Q, R \in \mathbb{K}[X]$ des polynômes tels que P soit premier avec Q. Alors

$$\operatorname{pgcd}(P, QR) = \operatorname{pgcd}(P, Q) \operatorname{pgcd}(P, R) \tag{5.68}$$

(2) En analogie avec le lemme 3.37, nous avons

$$\operatorname{pgcd}(P, PQ + R) = \operatorname{pgcd}(P, R). \tag{5.69}$$

Problèmes et choses à faire

Est-ce que ce lemme 5.43 est correct ? J'aimerais en voir une preuve.

5.3.4 Lemme et théorème de Gauss

Théorème 5.44 (Théorème de Gauss)**.**
Soient $P, Q, R \in \mathbb{K}[X]$ tels que P soit premier avec Q et divise QR. Alors P divise R.

Démonstration. Étant donné que P est premier avec Q, le théorème de Bézout[6] nous donne $U, V \in \mathbb{K}[X]$ tels que $PU + QV = 1$. De plus il existe un polynôme S tel que $PS = QR$. En multipliant l'identité de Bézout par R, nous obtenons

$$R = PUR + QVR = PUR + VPS = P(UR + VS), \tag{5.70}$$

ce qui signifie que P divise R. □

Le lemme suivant est une généralisation du lemme de Gauss dans $\mathbb{Z}$ (lemme 4.78).

5. Et non seulement deux à deux.
6. théorème 5.40.

Lemme 5.45 (Lemme de Gauss[34])**.**
Soient les polynômes unitaires $P, Q \in \mathbb{Q}[X]$. Si $PQ \in \mathbb{Z}[X]$, alors P et Q sont tous deux dans $\mathbb{Z}[X]$.

Démonstration. Soit $a > 0$ le plus petit entier tel que $aP \in \mathbb{Z}[X]$ (c'est le PPCM des dénominateurs) et de la même façon $b > 0$ le plus petit entier tel que $bQ \in \mathbb{Z}[X]$. On pose $P_1 = aP$ et $Q_1 = bQ$.

Si $ab = 1$, alors $a = b = 1$ et nous avons tout de suite $P, Q \in \mathbb{Z}[X]$. Nous supposons donc $ab > 1$ et nous considérons p, un diviseur premier de ab. Ensuite nous considérons la projection

$$\pi_p \colon \mathbb{Z}[X] \to (\mathbb{Z}/p\mathbb{Z})[X]. \tag{5.71}$$

Par définition $abPQ = P_1Q_1 \in \mathbb{Z}[X]$; en prenant la projection,

$$\pi_p(P_1)\pi_p(Q_1) = \pi_p(P_1Q_1) = \pi_P(ab)\pi_p(PQ) = 0 \tag{5.72}$$

parce que $\pi_p(ab) = 0$. Étant donné que $(\mathbb{Z}/p\mathbb{Z})[X]$ est intègre (exemple 4.41), nous avons soit $\pi_p(P_1) = 0$ soit $\pi_p(Q_1) = 0$. Supposons pour fixer les idées que $\pi_p(P_1) = 0$. Alors $P_1 = pP_2$ pour un certain $P_2 \in \mathbb{Z}[X]$. Par ailleurs P est unitaire et $P_1 = aP$, donc le coefficient de plus haut degré de P_1 est a, et nous concluons que p divise a.

Mettons $a = pa'$. Dans ce cas, $pa'P = P_1 = pP_2$, et donc $a'P = P_2 \in \mathbb{Z}[X]$. Cela contredit la minimalité de a. □

5.4 Extension de corps

Lemme 5.46.
Soit $\mathbb{L}$ un corps fini et $\mathbb{K}$ un sous corps de $\mathbb{K}$. Alors il existe $s \in \mathbb{N}$ tel que

$$\text{Card}(\mathbb{K}) = \text{Card}(\mathbb{L})^s. \tag{5.73}$$

Démonstration. Le corps $\mathbb{L}$ est un $\mathbb{K}$-espace vectoriel de dimension finie. Si s est la dimension alors nous avons la formule (5.73) parce que chaque élément de $\mathbb{L}$ est un s-uple d'éléments de $\mathbb{K}$. □

Définition 5.47.
*Soit $\mathbb{K}$ un corps commutatif. Une **extension** de $\mathbb{K}$ est un corps $\mathbb{L}$ muni d'un morphisme $i \colon \mathbb{K} \to \mathbb{L}$. Nous identifions le plus souvent $\mathbb{K}$ avec $i(\mathbb{K}) \subset \mathbb{L}$.*

*Une extension est **algébrique** de $\mathbb{K}$ est une extension dont tous les éléments sont racines de polynômes dans $\mathbb{K}[X]$.*

Notons que $\mathbb{R}$ n'est pas une extension algébrique de $\mathbb{Q}$. En effet il existe seulement une infinité *dénombrable* de polynômes dans $\mathbb{Q}[X]$ et donc une infinité dénombrable de racines de tels polynômes. Toute extension algébrique de $\mathbb{Q}$ est donc dénombrable.

5.4.1 Polynôme minimal

Lemme-définition 5.48 (Polynôme minimal)**.**
Soit $\mathbb{L}$ une extension de $\mathbb{K}$ et $a \in \mathbb{L}$. Nous considérons la partie

$$I_a = \{P \in \mathbb{K}[X] \text{ tel que } P(a) = 0\}. \tag{5.74}$$

(1) La partie I_a est un idéal principal dans $\mathbb{K}[X]$

(2) L'idéal I_a possède un générateur unitaire.

*Cet unique générateur unitaire est le **polynôme minimal** de a sur $\mathbb{K}$.*

Démonstration. Le fait que I_a soit un idéal est la définition du produit sur l'anneau des polynômes. Le théorème 5.37 fait le reste du travail : nous savons que $\mathbb{K}[X]$ est un anneau principal et donc que tous ses idéaux sont principaux (c'est la définition).

Le théorème 5.37 dit également que I_a possède un unique générateur unitaire. □

Exemple 5.49
Le polynôme minimal dépends du corps sur lequel on le considère. Par exemple le nombre imaginaire pur i accepte $X - i$ comme polynôme minimal sur $\mathbb{C}$ et $X^2 + 1$ sur $\mathbb{Q}[X]$. △

Proposition 5.50 ([1])**.**
Soit $\mathbb{L}$ un extension de $\mathbb{K}$ et $a \in \mathbb{L}$ dont le polynôme minimal sur $\mathbb{K}$ est $\mu_a \in \mathbb{K}[X]$. Alors

(1) le polynôme μ_a est irréductible[7] *sur $\mathbb{K}$;*

(2) si $Q \in \mathbb{K}[X]$ est un polynôme non annulateur de a alors Q et μ_a sont premiers[8] *entre eux.*

Démonstration. (1) Si μ_a était réductible sur $\mathbb{K}$, il existerait des polynômes P et Q dans $\mathbb{K}[X]$ tels que $\mu_a = PQ$; en évaluant cela en a nous obtenons $0 = P(a)Q(a)$, ce qui signifie que soit P soit Q est annulateur de a. Mais P et Q sont tous deux de degré supérieur à 1 (et donc de degré inférieur à celui de μ_a). Cela contredit la minimalité de μ_a : le polynôme minimal doit diviser tout polynôme annulateur.

(2) Soit Q un polynôme non annulateur de a. Si Q et μ_a ne sont pas premiers entre eux dans $\mathbb{K}[X]$ alors ils ont un diviseur commun. Mais μ_a étant irréductible sur $\mathbb{K}$ par (1), ce diviseur commun ne peut être que μ_a lui-même. Mais dans ce cas Q serait annulateur de a.

□

Deux éléments α et β dans $\mathbb{L}$ sont dit **conjugués** s'ils ont même polynôme minimal. Par exemple i et $-i$ sont conjugués dans $\mathbb{C}$ vu comme extension de $\mathbb{Q}$.

Lemme 5.51.
Un nombre complexe algébrique dont tous les conjugués sont de module 1 *est une racine de l'unité.*

Lemme 5.52.
Si $(\mathbb{L}, i)$ est une extension de $\mathbb{K}$, alors $\mathbb{L}$ est un espace vectoriel sur $\mathbb{K}$.

Démonstration. Il faut définir le produit d'un élément de $\mathbb{L}$ par un élément de $\mathbb{K}$; si $\lambda \in \mathbb{K}$ et $x \in \mathbb{L}$ nous la définissons par

$$\lambda \cdot x = i(\lambda)x \tag{5.75}$$

où la multiplication du membre de droite est celle du corps $\mathbb{L}$. □

Définition 5.53.
*Le **degré** de $\mathbb{L}$ est la dimension de cet espace vectoriel. Il est noté $[\mathbb{L} : \mathbb{K}]$; notons qu'il peut être infini.*

Exemple 5.54
L'ensemble $\mathbb{C}$ est une extension de $\mathbb{R}$ et son degré est $[\mathbb{C} : \mathbb{R}] = 2$. △

Proposition 5.55 ([42])**.**
Si $\mathbb{L}$ est une extension du corps $\mathbb{K}$ et si $\mathbb{M}$ est une extension de $\mathbb{L}$, alors les degrés se multiplient :

$$[\mathbb{M} : \mathbb{L}][\mathbb{L} : \mathbb{K}] = [\mathbb{M}; \mathbb{K}]. \tag{5.76}$$

Démonstration. Soit $\{l_i\}$ une base de $\mathbb{L}$ sur $\mathbb{K}$ et $\{m_j\}$ une base de $\mathbb{M}$ sur $\mathbb{L}$. Un élément de $\mathbb{M}$ se note

$$\sum_j b_j m_j \tag{5.77}$$

pour des $b_j \in \mathbb{L}$. Chacun de ces b_j s'écrit comme une combinaison des l_i :

$$\sum_j b_j m_j = \sum_{ij} (a_{ji} l_i) m_j, \tag{5.78}$$

7. Définition 5.29.
8. Définition 4.110.

donc nous voyons que les éléments $l_i m_j$ de $\mathbb{M}$ sont une base de $\mathbb{M}$ sur $\mathbb{K}$. □

Lemme 5.56.
Soit $\mathbb{L}$ un corps commutatif et $(\mathbb{K}_i)_{i\in I}$ une famille de sous-corps de $\mathbb{L}$. Alors $\bigcup_{i\in I}\mathbb{K}_i$ est un sous-corps de $\mathbb{L}$.

Définition 5.57.
Soit $\mathbb{L}$ une extension de $\mathbb{K}$ et $A\subset\mathbb{L}$.

(1) Nous notons $\mathbb{K}(A)$ le plus petit sous corps de $\mathbb{L}$ contenant $\mathbb{K}$ et A. C'est l'intersection de tous les sous-corps de $\mathbb{L}$ contenant A.

(2) Nous notons $\mathbb{K}[A]$ le plus petit sous anneau de $\mathbb{L}$ contenant $\mathbb{K}$ et A. C'est l'intersection de tous les sous-anneaux de $\mathbb{L}$ contenant A.

*Nous disons que l'extension $\mathbb{L}$ de $\mathbb{K}$ est **monogène** ou **simple** s'il existe $\theta\in\mathbb{L}$ tel que $\mathbb{L}=\mathbb{K}(\theta)$. Un tel élément θ est dit **élément primitif** de $\mathbb{L}$. Il n'est pas nécessairement unique.*

Remarque 5.58.
Les ensembles $\mathbb{K}(A)$ et $\mathbb{K}[A]$ sont aussi appelés corps et anneaux **engendré** par A. Cependant il faut bien remarquer que ce sont les parties de $\mathbb{L}$ engendrées par A. Il n'est pas question a priori de parler de corps engendré par A sans dire dans quel corps plus grand nous nous plaçons.

Exemple 5.59
Nous savons que $\mathbb{R}$ est une extension de $\mathbb{Q}$. Si $a\in\mathbb{R}$ alors $\mathbb{Q}(a)$ est le plus petit corps contenant $\mathbb{Q}$ et a. △

Exemple 5.60
Nous avons déjà vu à l'occasion de la définition 4.100 que $A[X]$ est l'anneau de tous les polynômes de degré fini en X. Cela rentre dans le cadre de la définition 5.57 parce un anneau contenant X doit contenir tous les X^n.

Notons que même si $\mathbb{K}$ est un corps, $\mathbb{K}[X]$ reste un anneau parce que A/X n'est pas dedans. Par contre, $\mathbb{K}(X)$ est un corps parce qu'il contient également les fractions rationnelles. △

Exemple 5.61
Si nous prenons $\mathbb{F}_5$ et que nous l'étendons par i, nous obtenons le corps $\mathbb{K}=\mathbb{F}_5(i)$. Nous savons que tous les éléments $a\in\mathbb{F}_5$ sont racines de X^5-X. Mais étant donné que $i^5=i$, nous avons aussi $x^5=x$ pour tout $x\in\mathbb{F}_5(i)$. Pour le prouver, utiliser le morphisme de Frobenius. Le polynôme X^5-X est donc le polynôme nul dans $\mathbb{K}$.

Ceci est un cas très particulier parce que nous avons étendus $\mathbb{F}_p$ par un élément α tel que $\alpha^p=\alpha$. En général sur $\mathbb{F}_p(\alpha)$, le polynôme X^p-X n'est pas identiquement nul, et possède donc au maximum p racines. Pour $x\in\mathbb{F}_p(\alpha)$, nous avons $x^p=x$ si et seulement si $x\in\mathbb{F}_p$. △

Lemme 5.62.
Soit $P\in\mathbb{K}[X]$ un polynôme unitaire irréductible de degré n. Il existe une extension $\mathbb{L}$ de $\mathbb{K}$ et $a\in\mathbb{L}$ telle que $\mathbb{L}=\mathbb{K}(a)$ et P est le polynôme minimal de a dans $\mathbb{L}$.

Démonstration. Nous prenons $\mathbb{L}=\mathbb{K}[X]/(P)$ où (P) est l'idéal dans $\mathbb{K}[X]$ généré par P. Cela est un corps par le corollaire 5.38. Nous identifions $\mathbb{K}$ avec $\phi(\mathbb{K})$ où

$$\phi\colon\mathbb{K}[X]\to\mathbb{L} \tag{5.79}$$

est la projection canonique. Nous considérons également $a=\phi(X)$.

Nous avons alors $P(a)=0$ dans $\mathbb{L}$. En effet $P(a)=P\big(\phi(X)\big)$ est à voir comme l'application du polynôme P au polynôme X, le résultat étant encore un élément de $\mathbb{L}$. En l'occurrence le résultat est P qui vaut 0 dans $\mathbb{L}$.

Le polynôme P étant unitaire et irréductible, il est minimum dans $\mathbb{L}$.

Nous devons encore montrer que $\mathbb{L} = \mathbb{K}(a)$. Le fait que $\mathbb{K}(a) \subset \mathbb{L}$ est une tautologie parce qu'on calcule $\mathbb{K}(a)$ dans $\mathbb{L}$. Pour l'inclusion inverse soit $Q(X) = \sum_i Q_i X^i$ dans $\mathbb{K}[X]$. Dans $\mathbb{L}$ nous avons évidemment $Q = \sum_i Q_i a^i$. □

Proposition 5.63.
Soit $\mathbb{K}$, un corps et $P \in \mathbb{K}[X]$ un polynôme. Soient a et b, deux racines de P dans (éventuellement) une extension $\mathbb{L}$ de $\mathbb{K}$. Si μ_a et μ_b sont les polynômes minimaux de a et b (dans $\mathbb{K}[X]$) et si $\mu_a \neq \mu_b$, alors $\mu_a\mu_b$ divise P dans $\mathbb{K}[X]$.

Démonstration. Nous considérons les idéaux

$$I_a = \{Q \in \mathbb{K}[X] \text{ tel que } Q(a) = 0\}; \tag{5.80a}$$
$$I_b = \{Q \in \mathbb{K}[X] \text{ tel que } Q(b) = 0\}; \tag{5.80b}$$

même si $Q(a)$ est calculé dans $\mathbb{L}$, ce sont des idéaux de $\mathbb{K}[X]$. Le polynôme μ_a est par définition le générateur unitaire de I_a, et vu que a est une racine de P, nous avons $P \in I_a$ et il existe un polynôme $Q \in \mathbb{K}[X]$ tel que

$$P = \mu_a Q. \tag{5.81}$$

Montrons que $\mu_a(b) \neq 0$. En effet supposons que $\mu_a(b) = 0$. Nous avons alors $\mu_a \in I_b$ et il existe $R \in \mathbb{K}[X]$ tel que $\mu_a = \mu_b R$. Étant donné que μ_b divise μ_a et que $\mu_b \neq \mu_a$, nous ne pouvons pas avoir $\mu_b(a) = 0$ parce que μ_b divise μ_a et que μ_a est minimal. Du coup nous devons avoir $R(a) = 0$, ce qui contredirait la minimalité de μ_a.

Étant donné que $\mu_a(b) \neq 0$, l'évaluation de (5.81) en b montre que $Q(b) = 0$, de telle sorte que $Q \in I_b$ et il existe une polynôme S tel que $Q = \mu_b S$, c'est à dire tel que $P = \mu_a\mu_b S$, ce qui signifie que $\mu_a\mu_b$ divise P. □

Exemple 5.64
Soit $P = (X^2 + 1)(X^2 + 2)$ dans $\mathbb{R}[X]$. Dans $\mathbb{C}$ nous avons les racines $a = i$ et $b = \sqrt{2}i$ dont les polynômes minimaux sont $\mu_a = X^2 + 1$ et $\mu_2 = X^2 + 2$. Nous avons effectivement $\mu_a\mu_b$ divise P dans $\mathbb{R}[X]$.

Si par contre nous considérions les racines $a = i$ et $b = -i$, nous aurions $\mu_a = \mu_n = X^2 + 1$, et les polynôme μ_a^2 ne divise pas P. △

5.4.2 Racines de polynômes

Proposition 5.65.
Soit un corps $\mathbb{K}$ et une extension $\mathbb{L}$. Soit $P \in \mathbb{K}[X]$ et $a \in \mathbb{L}$, une racine de P. Alors le polynôme minimal d'une racine divise[9] *tout polynôme annulateur.*

Autrement dit, l'idéal engendré par le polynôme minimal est l'idéal des polynômes annulateurs.

Démonstration. Nous considérons l'idéal

$$I = \{Q \in \mathbb{K}[X] \text{ tel que } Q(a) = 0\}. \tag{5.82}$$

Le fait que cela soit un idéal est simplement dû à la définition du produit : $(PQ)(a) = P(a)Q(a)$. Par le théorème 5.37, le polynôme minimal μ_a de a est dans I et qui plus est le génère : $I = (\mu_a)$. Par conséquent tout polynôme annulateur de a est divisé par μ_a. □

Corollaire 5.66 (Factorisation d'une racine)**.**
Soit $P \in \mathbb{K}[X]$, un polynôme de degré n et $\alpha \in \mathbb{K}$ tel que $P(\alpha) = 0$. Alors il existe un polynôme Q de degré $n - 1$ tel que $P(x) = (X - \alpha)Q$.

9. Définition 4.107.

Démonstration. Il s'agit d'un cas particulier de la proposition 5.65 : si $\alpha \in \mathbb{K}$ alors son polynôme minimal dans $\mathbb{K}$ est $X - \alpha$; donc $X - \alpha$ divise P. Il existe un polynôme Q tel que $P = (X - \alpha)Q$. Le degré est alors immédiat. □

Corollaire 5.67.
Soit $\mathbb{K}$ un corps commutatif et $\mathbb{L}$ une extension de $\mathbb{K}$. Si $P \in \mathbb{K}[X]$ est unitaire, annulateur et irréductible, alors il est le polynôme minimal[10] *de α sur $\mathbb{K}$.*

Démonstration. Par la proposition 5.65, le polynôme P est divisé par le polynôme minimal, mais P étant irréductible et unitaire c'est que P est le polynôme minimal. □

Théorème 5.68 (Polynôme qui a tellement de racines qu'il s'annule)**.**
Soit $\mathbb{K}$ un corps et $P \in \mathbb{K}[X]$ un polynôme de degré n possédant $n + 1$ racines distinctes $\alpha_1, \ldots, \alpha_{n+1}$, alors $P = 0$.

Voir la remarque 4.102 pour savoir ce qu'est un polynôme nul.

Démonstration. Si P est de degré 1, il s'écrit $P = aX + b$; s'il a comme racines α et β, nous avons le système

$$\begin{cases} a\alpha + b = 0 & \text{(5.83a)} \\ a\beta + b = 0. & \text{(5.83b)} \end{cases}$$

La différence entre les deux donne $a(\alpha - \beta) = 0$. Vu que $\alpha \neq \beta$, la règle du produit nul (lemme 4.38) nous donne $a = 0$. Maintenant que $a = 0$, l'annulation de b est alors immédiate.

Nous faisons maintenant la récurrence en supposant le théorème vrai pour le degré n et en considérant un polynôme P de degré $n + 1$ possédant $n + 2$ racines distinctes. Vu que $P(\alpha_1) = 0$, le corollaire 5.66 nous donne un polynôme Q de degré n tel que

$$P(X) = (X - \alpha_1)Q. \tag{5.84}$$

Étant donne que pour tout $i \neq 1$ nous avons $\alpha_i \neq \alpha_1$,

$$0 = P(\alpha_i) = \underbrace{(\alpha_i - \alpha_1)}_{\neq 0} Q(\alpha_i), \tag{5.85}$$

et la règle du produit nul donne $Q(\alpha_i) = 0$. Par conséquent le polynôme Q est de degré n et possède $n + 1$ racines distinctes ; tous ses coefficients sont alors nuls par hypothèse de récurrence. Tous les coefficients du produit (5.84) sont alors également nuls. □

Remarque 5.69.
Dans le cas des polynôme à plusieurs indéterminées, le résultat n'est plus vrai[11], mais nous avons la proposition 5.114.

Remarque 5.70.
L'intérêt du théorème 5.68 est que si l'on prouve qu'un polynôme s'annule sur un corps infini, alors il s'annulera sur n'importe quel autre corps. Nous aurons un exemple d'utilisation de cela dans le théorème de Cayley-Hamilton 12.339.

10. Définition 5.48.
11. Exemple 5.115.

5.4.3 Corps de rupture

Définition 5.71.
Soit $P \in \mathbb{K}[X]$ un polynôme irréductible. Une extension $\mathbb{L}$ de $\mathbb{K}$ est un ***corps de rupture*** *pour P s'il existe $a \in \mathbb{L}$ tel que $P(a) = 0$ et $\mathbb{L} = \mathbb{K}(a)$.*

Exemple 5.72
Soit $\mathbb{K} = \mathbb{Q}$ et $P(X) = X^2 - 2$. On pose $a = \sqrt{2}$ et $\mathbb{L} = \mathbb{Q}(\sqrt{2}) \subset \mathbb{R}$. De cette façon P est scindé :

$$P = (X - \sqrt{2})(X + \sqrt{2}). \tag{5.86}$$

Le corps $\mathbb{Q}(\sqrt{2})$ est donc un corps de rupture pour P. △

Exemple 5.73
Dans l'exemple 5.72, le polynôme P était scindé dans son corps de rupture. Il n'en est pas toujours ainsi. Prenons

$$P(X) = X^3 - 2 \tag{5.87}$$

et $a = \sqrt[3]{2}$. Nous avons, certes, $P(a) = 0$ dans $\mathbb{Q}(\sqrt[3]{2})$, mais P n'est pas scindé parce qu'il y a deux racines complexes. △

Exemple 5.74
Nous considérons le corps $\mathbb{Z}/p\mathbb{Z}$ où p est un nombre premier. Si $s \in \mathbb{Z}/p\mathbb{Z}$ n'est pas un carré, alors le polynôme $P = X^2 + s$ est irréductible et un corps de rupture de P sur $\mathbb{Z}/p\mathbb{Z}$ est donné par $(\mathbb{Z}/p\mathbb{Z})[X]/(X^2 + s)$, c'est à dire l'ensemble des polynômes de degrés 1 en $\sqrt{s}$. Le cardinal en est p^2. △

Proposition 5.75 (Propriétés d'extensions algébriques[1])**.**
Soit $\mathbb{K}$ un corps commutatif[12] *et a un élément algébrique sur $\mathbb{K}$, de polynôme minimal μ_a de degré n. Alors*

(1) En considérant l'application d'évaluation

$$\begin{aligned} \varphi_a \colon \mathbb{K}[X] &\to \mathbb{L} \\ Q &\mapsto Q(a), \end{aligned} \tag{5.88}$$

nous avons $\mathbb{K}[a] = \text{Image}(\varphi_a)$.

(2) Une base de $\mathbb{K}[a]$ comme espace vectoriel sur $\mathbb{K}$ est donnée par $\{1, a, a^2, \ldots, a^{n-1}\}$.

(3) Le degré de l'extension $\mathbb{K}[a]$ est égal au degré du polynôme minimal :

$$\big[\mathbb{K}[a] : \mathbb{K}\big] = n. \tag{5.89}$$

(4) L'anneau $\mathbb{K}[a]$ est l'ensemble des polynômes en a de degré $n - 1$ à coefficient dans $\mathbb{K}$.

(5) $\mathbb{K}(a) = \mathbb{K}[a]$.

(6) $\mathbb{K}[a] \simeq \mathbb{K}[X]/(\mu_a)$ (isomorphisme d'anneau).

L'intérêt de (6) est qu'il permet de caractériser $\mathbb{K}[a]$ sans avoir recours à un sur-corps de $\mathbb{K}$. Le point (3) indique que le degré d'une extension algébrique est égal au degré du polynôme minimal.

Démonstration. (1) Nous avons $\mathbb{K}[a] \subset \text{Image}(\varphi_a)$ parce que $\text{Image}(\varphi_a)$ est lui-même un un sous-anneau de $\mathbb{L}$ contenant $\mathbb{K}$ et a. Pour rappel, $\mathbb{K}[a]$ est l'intersection de tous les tels sous-anneaux.

L'inclusion inverse est le fait que si $Q \in \mathbb{K}[X]$ alors $Q(a) \in \mathbb{K}[a]$ parce que $\mathbb{K}[a]$ est un anneau et contient donc tous les a^n.

12. Juste en passant nous rappelons que tous les corps considérés ici sont commutatifs

(2) La partie $\{1, a, a^2, \ldots, a^{n-1}\}$ est libre parce qu'une combinaison linéaire de ces éléments est un polynôme de degré $n-1$ en a. Un tel polynôme ne peut pas être nul parce que nous avons mis comme hypothèse que le polynôme minimal de a est n.

Rappelons qu'en vertu de la définition 5.48, le polynôme minimal μ_a est unitaire ; donc le polynôme $\mu_a(X) - X^n$ est un polynôme de degré $n-1$. Par conséquent en posant $S(X) = X^n - \mu_a(X)$, le polynôme S est de degré $n-1$ et vérifie $a^n = S(a)$.

En vertu du point (1), un élément de $\mathbb{K}[a]$ s'écrit $Q(a)$ pour un certain $Q \in \mathbb{K}[X]$. Supposons que Q soit de degré $p > n-1$; alors nous le décomposons en une partie contenant les termes de degré jusqu'à $n-1$ et une partie contenant les autres :

$$Q(X) = Q_1(X) + X^n Q_2(X) \tag{5.90}$$

où Q_1 est de degré $n-1$ et Q_2 de degré $p-n$. Nous évaluons cette égalité en a :

$$Q(a) = Q_1(a) + S(a)Q_2(a). \tag{5.91}$$

Donc $Q(a)$ est l'image de a par le polynôme $Q_1 + SQ_2$ qui est de degré $p-1$. Par récurrence, $Q(a)$ est l'image de a par un polynôme de degré $n-1$.

Notons que l'idée est très simple : il s'agit de remplacer récursivement tous les a^n par $S(a)$.

(3) Conséquence immédiate de (2).

(4) Conséquence immédiate de (2).

(5) Un élément général non nul de $\mathbb{K}[a]$ est de la forme $Q(a)$ avec $Q \in \mathbb{K}[X]$; il s'agit de lui trouver un inverse. Pour cela nous remarquons que les polynômes $\mu_a(X)$ et $Q(x)$ sont premiers entre eux, sinon μ_a ne serait pas un polynôme minimal (voir la proposition 5.50). Donc le théorème de Bézout 5.40 affirme l'existence d'éléments $U, V \in \mathbb{K}[X]$ tels que

$$U\mu_a + VQ = 1 \tag{5.92}$$

dans $\mathbb{K}[X]$. Nous évaluons cette égalité en a en tenant compte de $\mu_a(a) = 0$ dans $\mathbb{K}[a]$:

$$U(a)\mu_a(a) + V(a)Q(a) = 1 \tag{5.93}$$

dans $\mathbb{K}[a]$. Par conséquent $V(a)Q(a) = 1$, ce qui signifie que $V(a)$ est l'inverse de $Q(a)$.

(6) Nous considérons l'application

$$\begin{aligned} \psi\colon \mathbb{K}[X]/(\mu_a) &\to \mathbb{K}[a] \\ \bar{R} &\mapsto R(a) \end{aligned} \tag{5.94}$$

et nous montrons qu'elle convient. Pour cela, nous nous souvenons que la proposition 5.65 nous enseigne que (μ_a), l'idéal engendré par μ_a, est égal à l'idéal des polynômes annulateurs de a dans $\mathbb{K}[X]$. Le polynôme μ_a divise tous les éléments de cet idéal ; voir aussi la définition 4.10 de l'idéal (μ_a). Cela étant mis au point, nous passons à la preuve.

ψ **est bien définie** Si $\bar{R} = \bar{S}$ alors $R = S + Q$ avec $Q \in (\mu_a)$, et par conséquent $R(a) = S(a) + Q(a)$ avec $Q(a) = 0$.

Surjective Nous savons que $\mathbb{K}[a] = \text{Image}(\varphi_a)$. Si $x \in \mathbb{K}[a]$ alors il existe $Q \in \mathbb{K}[X]$ tel que $x = Q(a)$. Dans ce cas nous avons aussi $x = \psi(\bar{Q})$.

Injective Si $\psi(\bar{R}) = 0$ alors $R(a) = 0$, mais comme mentionné plus haut, μ_a engendre l'idéal est polynômes annulateurs de a. Donc $R \in (\mu_a)$ et nous avons $\bar{R} = 0$ dans $\mathbb{K}[X]/(\mu_a)$.

□

Exemple 5.76

Un fait connu est que $\frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$. Donc l'inverse de $\sqrt{2}$ s'exprime bien comme un polynôme en $\sqrt{2}$ à

coefficients dans $\mathbb{Q}$, ce qui confirme le point (5) de la proposition 5.75. Du point de vue de Bézout, $\mu_{\sqrt{2}}(X) = X^2 - 2$, et nous cherchons des polynômes U et V tels que

$$U(X^2 - 2) + VX = 1. \tag{5.95}$$

cette égalité est réalisée par $U = -\frac{1}{2}$ et $V = \frac{1}{2}X$. Et effectivement $V(\sqrt{2})$ est bien l'inverse de $\sqrt{2}$:

$$V(\sqrt{(2)}) = \frac{1}{2}\sqrt{2}. \tag{5.96}$$

△

5.4.4 Corps de décomposition

Définition 5.77.
Soit $\mathbb{K}$ un corps commutatif et $F = (P_i)_{i\in I}$ une famille d'éléments non constants de $\mathbb{K}[X]$. Un ***corps de décomposition*** *de F est une extension $\mathbb{L}$ de $\mathbb{K}$ telle que*

(1) les P_i sont scindés sur $\mathbb{L}$,

(2) $\mathbb{L} = \mathbb{K}(R)$ où $R = \bigcup_{i\in I}\{x \in \mathbb{L}$ tel que $P_i(x) = 0\}$.

C'est à dire que $\mathbb{L}$ étends $\mathbb{K}$ par toutes les racines de tous les polynômes de F.

L'unicité est due à la proposition suivante.

Proposition 5.78.
Soit $\mathbb{K}$ un corps et $P \in \mathbb{K}[X]$. Soient $\mathbb{L}$ et $\mathbb{F}$ deux corps de décomposition de P. Alors il existe un isomorphisme $f\colon \mathbb{L} \to \mathbb{F}$ tel que $f|_{\mathbb{K}} = \mathrm{Id}$.

Nous pouvons donc parler du corps de décomposition d'un polynôme.

Soit $\mathbb{K}$, un corps et $\mathbb{L}$, une extension de $\mathbb{K}$. Un élément $a \in \mathbb{L}$ est **algébrique** sur $\mathbb{K}$ s'il existe un polynôme $P \in \mathbb{K}[X]$ tel que $P(a) = 0$.

Une **clôture algébrique** du corps $\mathbb{K}$ est une extension algébriquement close de $\mathbb{K}$ dont tous les éléments sont algébriques sur $\mathbb{K}$.

Remarque 5.79.
L'ensemble $\mathbb{C}$ n'est pas une clôture algébrique de $\mathbb{Q}$ parce qu'il existe des éléments de $\mathbb{C}$ qui ne sont pas des racines de polynômes à coefficients rationnels.

L'existence d'une clôture algébrique pour tout corps est le théorème de Steinitz.

Exemple 5.80
Soit p un nombre premier. Montrons que le polynôme

$$Q(X) = X^p - X + 1 \tag{5.97}$$

est irréductible dans $\mathbb{F}_p$.

Nous supposons qu'il n'est pas irréductible, c'est à dire que

$$Q(X) = R(X)S(X) \tag{5.98}$$

avec R et S, des polynômes de degrés $\geqslant 1$ dans $\mathbb{F}_p[X]$

Soit $\bar{\mathbb{F}}_p$ une clôture algébrique de $\mathbb{F}_p$ et $\alpha \in \bar{\mathbb{F}}_p$ tel que $R(\alpha) = 0$. Pour tout $a \in \mathbb{F}_p$, nous avons

$$\begin{aligned}
Q(\alpha + a) &= (\alpha + a)^p - (\alpha + a) + 1 && \text{(5.99a)}\\
&= \alpha^p + a^p - \alpha - a + 1 && \text{(5.99b)}\\
&= \alpha^p - \alpha + 1 && \text{(5.99c)}\\
&= Q(\alpha) && \text{(5.99d)}\\
&= 0 && \text{(5.99e)}
\end{aligned}$$

où nous avons utilisé le fait que $a^p = a$ et que α était une racine de Q. Ce que nous venons de prouver est que l'ensemble des racines de Q dans $\bar{\mathbb{F}}_p$ est donné par $\{\alpha + a \text{ tel que } a \in \mathbb{F}_p\}$.

Les polynômes R et S sont donc formés de produits de termes $X - (\alpha + a)$ avec $a \in \mathbb{F}_p$. L'un des deux –disons R pour fixer les idées– doit bien en avoir plus que 1. Nous avons alors

$$R(X) = \prod_{i=1}^{k} \big(X - (\alpha + a_i)\big) \tag{5.100}$$

où les a_i sont les éléments de $\mathbb{F}_p$. En développant un peu,

$$R(X) = X^k - \sum_{i=1}^{k}(\alpha + a_i^{k-1}) + \text{termes de degré plus bas en } X. \tag{5.101}$$

Le coefficient devant X^{k-1} n'est autre que $k\alpha + \sum_i a_i$. Étant donné que $k \neq 0$ et que $R \in \mathbb{F}_p[X]$, nous devons avoir $\alpha \in \mathbb{F}_p$. Par conséquent nous avons $\alpha^p = \alpha$ et une contradiction :

$$Q(\alpha) = \alpha^p - \alpha + 1 = 1 \neq 0. \tag{5.102}$$

Le polynôme $X^p - X + 1$ est donc irréductible sur $\mathbb{F}_p$. △

5.4.5 Clôture algébrique

Théorème 5.81.
Tout corps $\mathbb{K}$ possède une clôture algébrique Ω. De plus si $\mathbb{L}$ est une extension de $\mathbb{K}$, alors $\mathbb{L}$ est $\mathbb{K}$-isomorphe à un sous corps de Ω.

Les deux parties de ce théorème utilisent l'axiome du choix.

Notons en particulier que si Ω' est une autre clôture algébrique de $\mathbb{K}$, alors Ω et Ω' sont des sous corps l'un de l'autre et sont donc $\mathbb{K}$-isomorphes.

Lemme 5.82.
Les polynômes $P, Q \in \mathbb{K}[X]$ ne sont pas premiers entre eux si et seulement s'ils ont une racine commune dans la clôture algébrique Ω de $\mathbb{K}$.

Démonstration. Soit A un polynôme non inversible divisant P et Q. Par définition de Ω, ce polynôme A a une racine dans Ω qui est alors une racine commune à P et Q dans Ω.

Pour le sens inverse, si α est une racine commune de P et Q, alors le polynôme $X - \alpha$ divise P et Q et donc P et Q ne sont pas premiers entre eux. □

5.4.6 Extensions séparables

Notons que dans ce qui va suivre nous allons parler de $\mathbb{K}[X]$, l'ensemble des polynômes sur un corps. Cela ne s'applique donc pas à $\mathbb{Z}[X]$ par exemple.

Une des choses intéressantes avec les extensions séparables c'est qu'elles vérifient le théorème de l'élément primitif (5.95).

Définition 5.83.
Soit $\mathbb{K}$ un corps. Un polynôme irréductible *$P \in \mathbb{K}[X]$ est **séparable** sur $\mathbb{K}$ si dans un corps de décomposition, ses racines sont distinctes.*

*Si P est un polynôme non constant dont la décomposition en irréductibles est $P = P_1 \dots P_r$, nous disons qu'il est **séparable** si tous les P_i le sont.*

La proposition suivante donne un sens à la définition de polynôme irréductible séparable.

Proposition 5.84.
Soit P irréductible dans $\mathbb{K}[X]$ ayant des racines distinctes dans le corps de décomposition $\mathbb{L}$. Si $\mathbb{L}'$ est un autre corps de décomposition pour P, alors P a aussi ses racines distinctes dans $\mathbb{L}$.

Démonstration. L'ingrédient est la proposition 5.78 qui donne l'unicité du corps de décomposition à $\mathbb{K}$-isomorphisme près. Soit donc $\psi\colon \mathbb{L} \to \mathbb{L}'$ un isomorphisme laissant invariant les éléments de $\mathbb{K}$. D'une part, étant donné que P est à coefficients dans $\mathbb{K}$, nous avons $\psi(P) = P$. D'autre part dans $\mathbb{L}$ le polynôme P s'écrit

$$P = a(X - \alpha_1)\dots(X - \alpha_n) \tag{5.103}$$

avec $a \in \mathbb{K}$ et $\alpha_i \in \mathbb{L}$. Nous avons donc

$$P = \psi(P) = a(X - \psi(\alpha_1))\dots(X - \psi(\alpha_n)). \tag{5.104}$$

Donc les racines de P dans $\mathbb{L}'$ sont les éléments $\psi(\alpha_i)$ qui sont distincts. □

Exemple 5.85
Un polynôme peut être séparable sur un corps, mais non séparable sur un autre. Soit $\mathbb{L} = \mathbb{F}_p(T)$ et $\mathbb{K} = \mathbb{F}_p(T^p)$. Nous considérons le polynôme

$$P = X^p - T^p \tag{5.105}$$

dans $\mathbb{K}[X]$. Par le morphisme de Frobenius nous avons

$$P = (X - T)^p \tag{5.106}$$

dans $\mathbb{L}[X]$. Le polynôme P est irréductible sur $\mathbb{K}[X]$ parce que ses diviseurs sont de la forme $(X - T)^k$ qui contiennent T^k qui n'est pas dans $\mathbb{K}$ (sauf si $k = n$ ou $k = 0$).

Ce polynôme n'est pas séparable sur $\mathbb{K}$ parce que dans le corps de décomposition $\mathbb{L}$, la racine T est multiple. Notons bien le raisonnement : P étant irréductible, pour savoir s'il est séparable, on le regarde dans un corps de décomposition.

Par contre si nous regardons P dans $\mathbb{L}[X]$ alors P n'est plus irréductible parce que ses facteurs irréductibles sont $(X - T)$. N'étant pas irréductible, nous regardons les racines de *ses facteurs irréductibles*. Or chacun des facteurs irréductibles étant $X - T$, les racines sont simples. △

Exemple 5.86
Le polynôme $(X - 1)^3$ est séparable sur $\mathbb{Q}$ parce que ses facteurs irréductibles dans $\mathbb{Q}[X]$ sont $X - 1$ qui ont des racines simples. △

Exemple 5.87
Le polynôme $(X^2+1)^2$ est séparable dans $\mathbb{Q}[X]$. En effet, il a pour facteurs irréductible le polynôme $X^2 + 1$ dont les racines sont $\pm i$ dans l'extension $\mathbb{Q}(i)$. △

Proposition 5.88 ([43])**.**
Soit $P \in \mathbb{K}[X]$ un polynôme non constant. Les propriétés suivantes sont équivalentes.

1. *P a une racine multiple dans une extension de $\mathbb{K}$. C'est à dire qu'il existe une extension de $\mathbb{K}$ dans laquelle P a une racine multiple.*
2. *P a une racine multiple dans tout corps de décomposition .*
3. *P et P' ont une racine commune dans une extension de $\mathbb{K}$.*
4. *le degré de $\operatorname{pgcd}(P, P')$ est $\geqslant 1$.*

Démonstration. **(1)⇒(2)** Soit a, une racine multiple de P dans une extension $\mathbb{L}$ de $\mathbb{K}$, et $\mathbb{E}$, un corps de décomposition de P. Alors nous voulons prouver que P ait une racine multiple dans $\mathbb{E}$.

Nous pouvons voir $P \in \mathbb{L}[X]$, et construire une corps de décomposition $\mathbb{E}'$ qui est une extension de $\mathbb{L}$. Vu que $\mathbb{E}$ et $\mathbb{E}'$ sont deux corps de décomposition de P nous avons un isomorphisme $\psi\colon \mathbb{E} \to \mathbb{E}'$. Si $a \in \mathbb{E}$ est une racine multiple de P, alors $\psi(a)$ est une racine multiple de P dans $\mathbb{E}'$ parce que

$$P\big(\psi(a)\big) = \psi\big(P(a)\big). \tag{5.107}$$

(1)⇒(3) Soit $\mathbb{L}$ un corps de décomposition de P sur $\mathbb{K}$ et $a \in \mathbb{L}$, une racine multiple de P. On a alors $P = (X-a)^2Q$ avec $Q \in \mathbb{L}[X]$. En dérivant,

$$P' = 2(X-a)Q + (X-a)^2Q', \tag{5.108}$$

et donc a est également une racine de P'.

(3)⇒(4) Soit D un pgcd de P et P'. D'après le théorème de Bézout il existe $A, B \in \mathbb{K}[X]$ tels que

$$AP + BP' = D. \tag{5.109}$$

Si a est une racine commune de P et P' dans une extension $\mathbb{L}$, alors c'est aussi une racine de D et donc $\deg(D) \geqslant 1$.

(4)⇒(1) Si le degré de D est plus grand ou égal à 1, alors nous considérons une racine a de D dans $\mathbb{L}$ (une extension de $\mathbb{K}$). Étant donné que D divise P et P', l'élément a est une racine commune de P et P'. Nous montrons maintenant que a est alors une racine multiple de P. Vu que $P(a) = 0$ nous avons

$$P = (X-a)Q, \tag{5.110}$$

et $P' = Q + (X-a)Q'$. Mais alors $P'(a) = Q(a)$ et donc $Q(a) = 0$ et donc a est une racine double de P. Par conséquent a est une racine multiple de P dans $\mathbb{K}$.

□

Notons que si P est irréductible, cette proposition donne des conditions pour que P ne soit pas séparable.

Proposition 5.89.
Soit $P \in \mathbb{K}[X]$ irréductible. Le polynôme P est séparable si et seulement si $P' \neq 0$.

Démonstration. Soit $D = \text{pgcd}(P, P')$ et nous voudrions prouver que $\deg(D) \geqslant 1$ si et seulement si $P' = 0$. Si $P' = 0$, alors $\text{pgcd}(P, P') = P$ est donc $\deg'(D) \geqslant 1$.

Dans l'autre sens, si P est irréductible, il est associé à D parce qu'il n'a pas d'autres diviseurs que lui-même (à part 1). Nous avons donc $P = \lambda D$ avec $\lambda \in \mathbb{K}$ et donc $\deg(P) \geqslant 1$. Cela prouve immédiatement que $P' \neq 0$. □

Corollaire 5.90.
Si $\mathbb{K}$ est de caractéristique nulle, alors tout polynôme de $\mathbb{K}[X]$ est séparable.

Démonstration. Il suffit de montrer que les irréductibles sont séparables. Soit P un polynôme irréductible et unitaire de degré d. Le terme de plus haut degré de P' est alors dX^{d-1} qui est non nul parce que $d \neq 0$ en caractéristique nulle. Donc $P' \neq 0$ et donc P est séparable par la proposition 5.88. □

Définition 5.91.
Soit $\mathbb{L}$ une extension algébrique de $\mathbb{K}$.

*(1) On dit que l'élément $a \in \mathbb{L}$ est **séparable** sur $\mathbb{K}$ si son polynôme minimal dans $\mathbb{K}[X]$ est séparable sur $\mathbb{K}$.*

*(2) L'extension $\mathbb{L}$ est **séparable** si tous ses éléments sont séparables.*

Proposition 5.92.
Soit $\mathbb{K}$ un corps. Les conditions suivantes sont équivalentes :

(1) toutes les extensions algébriques de $\mathbb{K}$ sont séparables ;

(2) tout polynôme irréductible de $\mathbb{K}[X]$ est séparable.

En particulier les extensions algébriques des corps de caractéristique nulle sont toutes séparables.

Démonstration. Soit P un polynôme irréductible dans $\mathbb{K}[X]$. Soient a et b des racines de P dans un corps de décomposition $\mathbb{L}$. Par hypothèse, $\mathbb{L}$ est séparables, donc les polynômes minimaux μ_a et μ_b sont séparables dans $\mathbb{K}[X]$. Étant donné que P est irréductible, il est le seul à se diviser lui-même et donc $\mu_a = \mu_b = P$. Donc P est séparable sur $\mathbb{K}$.

Dans l'autre sens, soit $\mathbb{L}$ une extension algébrique de $\mathbb{K}$, soit $a \in \mathbb{L}$ et le polynôme minimal $\mu_a \in \mathbb{K}[X]$. Par définition il est irréductible et donc séparable (par hypothèse). Donc a est séparable et $\mathbb{L}$ est une extension séparable.

La dernière phrase est une conséquence du corollaire 5.90. □

Exemple 5.93
Une des conséquences les plus intéressantes de la proposition 5.92 est que toutes les extensions algébriques de $\mathbb{Q}$ sont séparables. △

Théorème 5.94 ([21])**.**
Soit $\mathbb{K}$ un corps (pas spécialement fini). Tout sous-groupe fini de $\mathbb{K}^*$ est cyclique.

Démonstration. Soit G un sous-groupe fini de $\mathbb{K}^*$ et ω son exposant (qui est le PPCM des ordres des éléments de G). Étant donné que $|G|$ est divisé par tous les ordres, il est divisé par le PPCM des ordres. Bref, nous avons

$$x^\omega = 1 \tag{5.111}$$

pour tout $x \in G$. Mais ce polynôme possède au plus ω racines dans $\mathbb{K}$. Du coup $|G| \leqslant \omega$. Et comme on avait déjà vu que $\omega \mid |G|$, on a $\omega = |G|$. Il suffit plus que trouver un élément d'ordre effectivement ω. Cela est fait par le lemme 3.58. □

Théorème 5.95 (Théorème de l'élément primitif[21])**.**
Toute extension de corps séparable finie admet un élément primitif.

Plus explicitement, soient $\alpha_1, \ldots, \alpha_n$ des éléments algébriques séparables sur $\mathbb{K}$; alors $\mathbb{L} = \mathbb{K}(\alpha_1, \ldots, \alpha_n)$ admet un élément primitif.

Démonstration. Si le corps $\mathbb{K}$ est fini, alors $\mathbb{L}$ est également fini. Donc $\mathbb{L}^*$ est cyclique par le théorème 5.94. Si θ est un générateur de $\mathbb{L}^*$, alors $\mathbb{L} = \mathbb{K}(\theta)$.

Passons au cas où $\mathbb{K}$ est infini. Il suffit d'examiner le cas $n = 2$; en effet pour $n = 1$ c'est trivial et si $n > 2$, alors

$$\mathbb{K}(\alpha_1, \ldots, \alpha_n) = \mathbb{K}(\alpha_1, \ldots, \alpha_{n-1})(\alpha_n), \tag{5.112}$$

et donc si $\mathbb{K}(\alpha_1, \ldots, \alpha_{n-1}) = \mathbb{K}(\theta)$, nous avons

$$\mathbb{K}(\alpha_1, \ldots, \alpha_n) = \mathbb{K}(\theta, \alpha_n) \tag{5.113}$$

et nous sommes réduit au cas $n = 2$ par récurrence.

Soit donc $\mathbb{L} = \mathbb{K}(\alpha, \beta)$; soit P le polynôme minimal de α sur $\mathbb{K}$ et Q celui de β. Nous nommons $\mathbb{E}$, un corps de décomposition de PQ. Nous avons $\mathbb{L} \subset \mathbb{E}$. Vu que P et Q sont polynômes minimaux d'éléments qui sont par hypothèse séparables, les polynômes P et Q sont séparables. Donc dans $\mathbb{E}$ les racines de P sont distinctes parce que P est irréductible (et idem pour Q). Soient les racines

$$\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_r \tag{5.114}$$

de P dans $\mathbb{E}$ et les racines

$$\beta_1 = \beta, \beta_2, \ldots, \beta_s \tag{5.115}$$

de Q dans $\mathbb{E}$. Ici r et s sont les degrés de P et Q.

Si $s = 1$ alors $Q = X - \beta$ et donc $\beta \in \mathbb{K}$ (parce que $Q \in \mathbb{K}[X]$). Du coup nous avons $\mathbb{L} = \mathbb{K}(\alpha)$ et le théorème est démontré. Nous supposons donc maintenant que $s \geqslant 2$.

Pour chaque $(i,j) \in [\![1,r]\!] \times [\![2,s]\!]$, l'équation $\alpha_i + x\beta_k = \alpha_1 + x\beta_1$ pour $x \in \mathbb{K}$ a au plus[13] une solution donnée le cas échéant par

$$x = (\alpha_i - \alpha_1)(\beta_1 - \beta_k)^{-1} \tag{5.116}$$

Notons que cela est de toutes façons dans $\mathbb{L}$ et qu'étant donné que $\beta_1 \neq \beta_k$, cette solution a un sens (ici on utilise l'hypothèse de séparabilité). Étant donné que $\mathbb{K}$ est infini nous pouvons donc trouver un $c \in \mathbb{K}$ qui ne résout aucune des équations (5.116) :

$$\alpha_i + c\beta_k \neq \alpha_1 + c\beta_1. \tag{5.117}$$

Nous posons $\theta = \alpha + c\beta$ et nous prétendons que $\mathbb{L} = \mathbb{K}(\theta)$. Montrons que $\beta = \mathbb{K}(\theta)$. Soit dans $\mathbb{K}(\theta)[T]$ les polynômes $Q(T)$ et $S(T) = P(\theta - cT)$. Nous nommons R le PGCD de ces deux polynômes.

D'une part, une racine de R doit être une racine de Q, et donc être un des β_i. D'autre part, le choix de c fait que β est une racine de R parce que

$$S(\beta) = P(\alpha + c\beta - c\beta) = P(\alpha) = 0. \tag{5.118}$$

Enfin si $k \geqslant 2$, alors

$$S(\beta_k) = P\big(\alpha_1 + c(\beta - \beta_k)\big) \neq 0 \tag{5.119}$$

parce que $\alpha_1 + c(\beta + \beta_k)$ n'est aucun des α_i. Nous concluons que β est l'unique racine de R et donc que

$$R = X - \beta \in \mathbb{K}(\theta)[T], \tag{5.120}$$

et donc $\beta \in \mathbb{K}(\theta)$.

De plus $\alpha = \theta - c\beta$ est alors immédiatement dans $\mathbb{K}(\theta)$. À partir du moment où α et β sont dans $\mathbb{K}(\theta)$, nous avons obtenu $\mathbb{L} = \mathbb{K}(\alpha,\beta) = \mathbb{K}(\theta)$.

□

Exemple 5.96
Le théorème de l'élément primitif 5.95 ne tient pas pour les corps non commutatifs. Par exemple si nous considérons pour $\mathbb{K}$ le corps des quaternions et comme G le groupe à 8 éléments $\{\pm 1, \pm i, \pm j, \pm k\}$. Ce dernier groupe n'est pas cyclique alors qu'il est un groupe fini dans $\mathbb{K}^*$. △

Exemple 5.97
Il est aussi possible pour un groupe fini d'avoir $\omega(G) = |G|$ sans pour autant que G soit cyclique. Par exemple pour $G = S_3$, nous avons $|S_3| = 6$ alors que les éléments de S_3 sont soit d'ordre 2 soit d'ordre 3 et $\omega(G) = \text{ppcm}(2,3) = 6$. Pourtant S_3 n'est pas cyclique. △

5.4.7 Idéal maximum

Définition 5.98.
*Un nombre (dans $\mathbb{C}$) est **transcendant** s'il n'est racine d'aucun polynôme non nul à coefficients entiers. Plus généralement si $\mathbb{L}$ est un extension du corps $\mathbb{K}$ alors si $t \in \mathbb{L}$ est une racine d'un polynôme dans $\mathbb{K}[X]$ nous disons que t est **algébrique** sur $\mathbb{K}$; sinon nous disons que t est **transcendant** sur $\mathbb{K}$.*

Définition 5.99.
*Une $\mathbb{K}$-algèbre est de **type fini** si elle est le quotient de $\mathbb{K}[X_1,\dots,X_n]$ par un idéal (pour un certain n).*

13. La solution (5.116) peut être dans $\mathbb{L}$ et non dans $\mathbb{K}$. L'équation peut donc très bien ne pas avoir de solutions $x \in \mathbb{K}$.

Théorème 5.100 (wikipédia)**.**
un idéal I d'un anneau commutatif $\mathbb{A}$ est maximal si et seulement si le quotient $\mathbb{A}/I$ est un corps.

Théorème 5.101 ([44])**.**
Soit $\mathbb{K}$ un corps et B, une $\mathbb{K}$-algèbre de type fini. Si B est un corps, alors c'est une extension algébrique finie de $\mathbb{K}$.

Théorème 5.102 ([44])**.**
Si $\mathbb{K}$ est un corps algébriquement clos, les idéaux maximaux de $\mathbb{K}[X_1, \dots, X_n]$ sont de la forme

$$(X_1 - a_1, \dots, X_n - a_n) \tag{5.121}$$

où les a_i sont des éléments de $\mathbb{K}$.

Démonstration. Nous commençons par montrer que

$$J = (X_1 - a_1, \dots, X_n - a_n) \tag{5.122}$$

est un idéal maximum. Pour cela nous considérons le morphisme surjectif d'anneaux

$$\begin{aligned} \phi\colon \mathbb{K}[X_1, \dots, X_n] &\to \mathbb{K} \\ P &\mapsto P(a_1, \dots, a_n). \end{aligned} \tag{5.123}$$

Soit $P \in \ker(\phi)$; nous écrivons la division euclidienne de P par $X - a_1$ puis celle du reste par $X - a_2$ et ainsi de suite :

$$P = (X - a_1)Q_1 + \cdots + (X_n - a_n)Q_n + R \tag{5.124}$$

où R doit être une constante parce que le premier reste est de degré zéro en X_1, le second est de degré zéro en X_1 et X_2, etc. Afin d'identifier cette constante, nous appliquons l'égalité (5.124) à $(a_1, \dots, a_n)$ et en nous rappelant que $P \in \ker(\phi)$ nous obtenons

$$0 = P(a_1, \dots, a_n) = R, \tag{5.125}$$

donc $R = 0$ et $P = (X_1 - a_1)Q_1 + \cdots + (X_n - a_n)Q_n$, c'est à dire $P \in J$. Nous avons donc $\ker(\phi) \subset J$. Par ailleurs $J \subset \ker(\phi)$ est évident, donc $J = \ker(\phi)$.

Vu que J est le noyau de l'application $\mathbb{K}[X_1, \dots, X_n] \to \mathbb{K}$, nous avons

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{J} = \mathbb{K}. \tag{5.126}$$

Donc J est un idéal maximal parce que tout polynôme n'étant pas dans J doit avoir un terme indépendant non nul et donc être dans $\mathbb{K}$ vis à vis du quotient $\mathbb{K}[X_1, \dots, X_n]/J$.

Nous montrons maintenant l'implication inverse. Nous supposons que I est un idéal maximum et nous montrons qu'il doit être égal à J (pour un certain choix de $a_1, \dots, a_n$).

Le quotient

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{I} \tag{5.127}$$

est une $\mathbb{K}$-algèbre de type fini (définition 5.99). De plus c'est un corps par le théorème 5.100. C'est donc une extension algébrique finie de $\mathbb{K}$ par le théorème 5.101. Mais $\mathbb{K}$ étant algébriquement clos, il est sa propre et unique extension algébrique ; nous en déduisons que

$$\frac{\mathbb{K}[X_1, \dots, X_n]}{I} = \mathbb{K}. \tag{5.128}$$

Donc pour tout $1 \leqslant i \leqslant n$, il existe $a_i \in \mathbb{K}$ tel que $X_i - a_i \in I$, sinon le monôme X_i ne se projetterait pas sur un élément dans $\mathbb{K}$ dans le quotient. Cela prouve que J est contenu dans I ; par maximalité nous avons donc $I = J$. □

Corollaire 5.103.
Soit $\mathbb{K}$ un corps algébriquement clos et I, un idéal de $\mathbb{K}[X_1, \dots, X_n]$. Si nous notons

$$V(I) = \{x \in \mathbb{K}^n \text{ tel que } P(x_1, \dots, x_n) = 0\} \tag{5.129}$$

l'ensemble des racines communes à tous les éléments de I, on a $V(I) = \varnothing$ si et seulement si $I = \mathbb{K}[X_1, \dots, X_n]$.

Démonstration. Si $I = \mathbb{K}[X_1, \dots, X_n]$ en particulier $1 \in I$ et nous avons évidemment $V(I) = \varnothing$. Le sens difficile est l'autre sens.

Supposons que $I \neq \mathbb{K}[X_1, \dots, X_n]$ et que K est un idéal maximum contenu dans I. Nous savons déjà par le théorème 5.102 que K est de la forme $K = (X_1 - a_1, \dots, X_n - a_n)$. Un élément de I est dans K, donc si $P \in I$ nous avons

$$P(a_1, \dots, a_n) = 0, \tag{5.130}$$

c'est à dire que $(a_1, \dots, a_n) \in V(I)$ et donc que $V(I) \neq 0$. □

5.5 Espaces de polynômes

Dans cette section nous abandonnons pour quelques minutes l'espace $\mathbb{R}^m$ et considérons plus attentivement l'espace des fonctions polynômiales $\mathcal{P}_{\mathbb{R}}$ et de ses sous-espaces $\mathcal{P}^k_{\mathbb{R}}$, pour k dans $\mathbb{N}_0$.

Attention : les polynômes en soi sont définis par la définition 4.100.

Pour chaque $k > 0$ donné nous définissons

$$\mathcal{P}^k_{\mathbb{R}} = \{p : \mathbb{R} \to \mathbb{R} \,|\, p : x \mapsto a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k,\ a_i \in \mathbb{R},\ \forall i = 0, \dots, k\}. \tag{5.131}$$

Il est facile de se convaincre que la somme de deux polynômes de degré inférieur ou égal à k est encore un polynôme de degré inférieur ou égal à k. En outre il est clair que la multiplication par un scalaire ne peut pas augmenter le degré d'un polynôme. L'ensemble $\mathcal{P}^k_{\mathbb{R}}$ est donc un espace vectoriel muni des opérations héritées de $\mathcal{P}_{\mathbb{R}}$.

La base canonique de l'espace $\mathcal{P}^k_{\mathbb{R}}$ est donné par les monômes $\mathcal{B} = \{x \mapsto x^j \,|\, j = 0, \dots, k\}$. Le fait que cela soit une base est vraiment facile à démontrer et est un exercice très utile si vous ne l'avez pas encore vu dans un cours précédent.

Nous allons maintenant étudier trois application linéaires de $\mathcal{P}^k_{\mathbb{R}}$ vers des autres espaces vectoriels

L'isomorphisme canonique $\phi : \mathcal{P}^k_{\mathbb{R}} \to \mathbb{R}^{k+1}$ Nous définissons ϕ par les relations suivantes

$$\phi(x^j) = e_{j+1}, \qquad \forall j \in \{0, \dots, k\}.$$

Cela veut dire que pour tout p dans $\mathcal{P}^k_{\mathbb{R}}$, avec $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^K$, l'image de p par ϕ est

$$\phi(p) = \phi\left(\sum_{j=0}^{k} a_j x^j\right) = \sum_{j=0}^{k} a_j e_{j+1}.$$

Exemple 5.104

Soit $k = 5$ on a

$$\phi(-8 - 7x - 4x^2 + 4x^3 + 2x^5) = \begin{pmatrix} -8 \\ -7 \\ -4 \\ 4 \\ 0 \\ 2 \end{pmatrix}. \tag{5.132}$$

△

Cette application est clairement bijective et respecte les opérations d'espace vectoriel, donc elle est un isomorphisme d'espaces vectoriels. L'existence d'un isomorphisme entre $\mathcal{P}^k_{\mathbb{R}}$ et $\mathbb{R}^{k+1}$ est un cas particulier du théorème qui dit que pour chaque m dans $\mathbb{N}_0$ fixée, tous les espaces vectoriels sur $\mathbb{R}$ de dimension m sont isomorphes à $\mathbb{R}^m$. Vous connaissez peut être déjà ce théorème depuis votre cours d'algèbre linéaire.

La dérivation $d : \mathcal{P}^k_{\mathbb{R}} \to \mathcal{P}^{k-1}_{\mathbb{R}}$ L'application de dérivation d fait exactement ce qu'on s'attend d'elle

$$d(x^0) = d(1) = 0, \qquad d(x^j) = jx^{j-1}, \quad \forall j \in \{1, \dots, k\}.$$

Cette application n'est pas injective, parce que l'image de p ne dépend pas de la valeur de a_0, donc si deux polynômes sont les mêmes à une constante près ils auront la même image par d.

Exemple 5.105
Soit $k = 3$ on a

$$d(-8 - 12x + 4x^3) = -12(1) + 4(3x^2) = -12 + 12x^2. \tag{5.133}$$

Noter que $d(-30 - 12x + 4x^3) = d(-8 - 12x + 4x^3)$. Cela confirme, comme mentionné plus haut, que la dérivée n'est pas injective. △

L'intégration $I : \mathcal{P}^k_{\mathbb{R}} \to \mathcal{P}^{k+1}_{\mathbb{R}}$ Nous pouvons définir une application que est «à une constante prés» l'application inverse de la dérivation

$$I(p) = \int_0^x p(t)\, dt. \tag{5.134}$$

Il faut comprendre que dans l'intégral la variable t est simplement la variable d'intégration. La «vraie» variable de la fonction image de p sera x !

Comme d'habitude nous écrivons explicitement l'action de I sur les éléments de la base canonique

$$I(x^j) = \int_0^x t^k\, dt = \frac{x^{j+1}}{j+1}. \tag{5.135}$$

Exemple 5.106
Soit $k = 4$ on a

$$I(6 + 2x + x^2 + x^4) = 6x + x^2 + \frac{x^3}{3} + \frac{x^5}{5}. \tag{5.136}$$

△

Remarquez que, étant donné que dans la définition de I nous avons décidé d'intégrer entre zéro et x, tous les polynômes dans $\mathcal{P}^{k+1}_{\mathbb{R}}$ qui sont l'image par I d'un polynôme de $\mathcal{P}^k_{\mathbb{R}}$ ont $a_0 = 0$. Cela veut dire que nous pouvons générer toute l'image de I en utilisant un sous-ensemble de la base canonique de $\mathcal{P}^{k+1}_{\mathbb{R}}$, en particulier $\mathcal{B}_1 = \{x \mapsto x^j \mid j = 1, \dots, k\} \subset \mathcal{B}$ nous suffira. Cela n'est guère surprenant, parce que l'image par une application linéaire d'un espace vectoriel de dimension finie ne peut pas être un espace de dimension supérieure.

Les applications de dérivation et intégration correspondent évidemment à des application linéaires de $\mathcal{P}_{\mathbb{R}}$ dans lui-même.

L'espace de tous les polynômes étant de dimension infinie, il peut servir de contre exemple assez simple. Dans la sous-section 10.1.2, nous verrons que toutes les normes ne sont pas équivalentes sur l'espace des polynômes.

5.5.1 Polynômes symétriques, alternés ou semi-symétriques

[34].

Soit $\mathbb{K}$ un corps de caractéristique différente[14] de 2. Le groupe S_n agit sur l'anneau $\mathbb{K}[T_1, \dots, T_n]$ par

$$(\sigma \cdot f)(T_1, \dots, T_n) = f\big(T_{\sigma(1)}, \dots, T_{\sigma(n)}\big). \tag{5.137}$$

On peut vérifier que c'est un action.

Définition 5.107.
Un polynôme Q en n indéterminées est

(1) ***symétrique*** *si $Q = \sigma \cdot Q$ pour tout $\sigma \in S_n$;*

(2) ***alterné*** *si $\sigma \cdot Q = \epsilon(\sigma)Q$ pour tout $\sigma \in S_n$;*

(3) ***semi-symétrique*** *si $\sigma \cdot Q = Q$ pour tout $\sigma \in A_n$*

Le polynôme $T_1 + T_2$ est symétrique ; le polynôme $T_1 + T_2^2$ ne l'est pas.

5.5.2 Polynôme symétrique élémentaire

Définition 5.108.
Le kième ***polynôme symétrique élémentaire*** *à n inconnues est le polynôme est*

$$\sigma_k(T_1, \dots, T_n) = \sum_{s \in F_k} \prod_{i=1}^{k} T_{s(i)} \tag{5.138}$$

où F_k est l'ensemble des fonctions strictement croissantes $\{1, 2, \dots, k\} \to \{1, 2, \dots, n\}$.

Une autre façon de décrire ces polynômes élémentaires est

$$\sigma_k = \sum_{1 \leqslant i_1 < \dots < i_k \leqslant n} X_{i_1} \dots X_{i_k}. \tag{5.139}$$

Par exemple

$$\sigma_1(T_1, \dots, T_n) = T_1 + T_2 + \dots + T_n \tag{5.140a}$$
$$\sigma_2(T_1, \dots, T_n) = T_1T_2 + \dots + T_1T_n + T_2T_3 + \dots + T_2T_n + \dots + T_{n-1}T_n \tag{5.140b}$$
$$\sigma_n(T_1, \dots, T_n) = T_1 \dots T_n. \tag{5.140c}$$

En particulier, $\sigma_2(x, y, z) = xy + yz + xz$.

Théorème 5.109 ([45])**.**
Si Q est un polynôme symétrique en $T_1, \dots, T_n$, alors il existe un et un seul polynôme P en n indéterminées tel que

$$Q(T_1, \dots, T_n) = P\big(\sigma_1(T_1, \dots, T_n), \dots, \sigma_n(T_1, \dots, T_n)\big). \tag{5.141}$$

Exemple 5.110
Nous voulons décomposer $P(x, y, z) = x^3 + y^3 + z^3$ en polynômes symétriques élémentaires, c'est à dire en

$$\begin{cases} \sigma_1 = x + y + z & (5.142a) \\ \sigma_2 = xy + xz + yz & (5.142b) \\ \sigma_3 = xyz. & (5.142c) \end{cases}$$

Étant donné que P est de degré 3, les seules combinaisons des σ_i qui peuvent intervenir sont σ_1^3, $\sigma_1\sigma_2$ et σ_3. Étant donné que dans P le coefficient de x^3 est un, il est obligatoire d'avoir un coefficient 1 devant σ_1^3. Nous le calculons :

14. Le truc de la caractéristique deux est que $a = -a$ n'implique pas $a = 0$.

```
----------------------------------------------------------------------
| Sage Version 4.8, Release Date: 2012-01-20                         |
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: var('x,y,z')
(x, y, z)
sage: P=x**3+y**3+z**3
sage: S1=x+y+z
sage: S2=x*y+x*z+y*z
sage: S3=x*y*z
sage: (S1**3).expand()
x^3 + 3*x^2*y + 3*x^2*z + 3*x*y^2 + 6*x*y*z + 3*x*z^2 + y^3 + 3*y^2*z + 3*y*z^2 + z^3
sage: (S1**3-P).expand()
3*x^2*y + 3*x^2*z + 3*x*y^2 + 6*x*y*z + 3*x*z^2 + 3*y^2*z + 3*y*z^2
x^3 + 3*x^2*y + 3*x^2*z + 3*x*y^2 + 6*x*y*z + 3*x*z^2 + y^3 + 3*y^2*z + 3*y*z^2 + z^3
```

Dans la différence $\sigma_1^3 - P$ nous voyons que le terme en xyz est $6xyz$; par conséquent nous savons que le coefficient de σ_3 sera -6. Il nous reste :

```
sage: (S1**3+6*S3-P).expand()
3*x^2*y + 3*x^2*z + 3*x*y^2 + 12*x*y*z + 3*x*z^2 + 3*y^2*z + 3*y*z^2
```

que nous identifions facilement avec $3\sigma_1\sigma_2$. Nous avons donc

$$P = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3. \tag{5.143}$$

△

Lemme 5.111 ([34])**.**
Soit $\mathbb{K}$ une extension de degré δ de $\mathbb{Q}$ et $P \in \mathbb{K}[T_1, \ldots, T_m]$. Alors il existe $\bar{P} \in \mathbb{Q}[T_1, \ldots, T_m]$ tel que

(1) $\deg \bar{P} = \delta \deg(P)$

(2) pour tout $(z_1, \ldots, z_m) \in \mathbb{C}^m$ tel que $P(z_1, \ldots, z_m) = 0$, on a $\bar{P}(z_1, \ldots, z_m) = 0$.

Démonstration. En vertu de la proposition 5.92 et de l'exemple 5.93, $\mathbb{K}$ est une extension séparable de $\mathbb{Q}$, et donc vérifie le théorème de l'élément primitif (5.95). Il existe $\theta \in \mathbb{K}$ tel que $\mathbb{K} = \mathbb{Q}(\theta)$. Soit $P_\theta \in \mathbb{Q}[X]$ le polynôme minimal de θ. L'extension $\mathbb{K}$ étant de degré δ, et θ étant un générateur, une base de $\mathbb{K}$ comme espace vectoriel sur $\mathbb{Q}$ est

$$\{1, \theta, \ldots, \theta^{\delta-1}\}. \tag{5.144}$$

Mais par ailleurs la proposition 5.75(2) nous indique qu'une base de $\mathbb{Q}(\theta)$ sur $\mathbb{Q}$ est donnée par

$$\{1, \theta, \ldots, \theta^{n-1}\} \tag{5.145}$$

où n est le degré de P_θ. Donc P_θ est de degré δ. Nous nommons $\theta_1, \ldots, \theta_\delta$ les racines de P_θ dans un corps de décomposition. Ici nous notons $\theta = \theta_1$ et nous ne prétendons pas que $\theta_k \in \mathbb{K}$. Notons que ces θ_i sont toutes des racines simples de P_θ, sinon nous aurions un facteur irréductible $(X - \theta_k)^2$, et P_θ ne serait pas irréductible sur $\mathbb{Q}$.

Soit σ_k le morphisme canonique

$$\begin{aligned} \sigma_k \colon \mathbb{Q}(\theta) &\to \mathbb{Q}(\theta_k) \\ \sum_i q_i\theta^i &\mapsto \sum_i q_i\theta_k^i \end{aligned} \tag{5.146}$$

Nous avons $\sigma_1 : \mathbb{K} \to \mathbb{K}$ qui est l'identité.

Notons N le degré du polynôme $P \in \mathbb{K}[T_1, \dots, T_m]$ dont il est question dans l'énoncé. Nous le décomposons alors en

$$P = \sum_{l=0}^{N} \sum_{i=1}^{m} c_{il} T_i^l \tag{5.147}$$

avec $c_{il} \in \mathbb{K}$. Nous voyons $c_{i,\cdot}$ comme un élément de $\mathbb{K}^m$ et donc nous écrivons [15]

$$P = \sum_{l=0}^{N} \sum_{i=1}^{m} c_l(\theta)_i T_i^l \tag{5.148}$$

où $c_l \in \mathbb{Q}[X]^m$. Nous pouvons choisir $\deg(c_l) < \delta$ parce que les puissances plus grandes de θ ne génèrent rien de nouveau.

Nous posons aussi

$$P^{\sigma_k} = \sum_{l,i} c_l(\theta_k)_i T_i^l \in \mathbb{Q}(\theta_k)[T_1, \dots, T_m], \tag{5.149}$$

et $\bar{P} = P P^{\sigma_2} \dots P^{\sigma_k}$. Le coefficient de T_i^l dans $\bar{P}$ est

$$\bar{c}_l(\theta_1, \dots, \theta_\delta)_i = \sum_{l_1 + \dots + l_\delta = l} c_{l_1}(\theta_1)_i \dots c_{l_\delta}(\theta_\delta)_i. \tag{5.150}$$

Ce dernier est un polynôme en les θ_k à coefficients dans $\mathbb{Q}$. Qui plus est, c'est un polynôme symétrique. En effet un terme contenant $\theta_k^a \theta_l^b$ provenant de $c_{l_i}(\theta_k) c_{l_j}(\theta_l)$ a un terme correspondant $\theta_k^b \theta_l^a$ provenant de $c_{l_j}(\theta_k) c_{l_i}(\theta_l)$.

C'est donc le moment d'utiliser le théorème 5.109 à propos des polynômes symétriques élémentaires qui nous dit que les coefficients de $\bar{P}$ sont en réalité des polynômes en ceux de P_θ qui sont dans $\mathbb{Q}$. Donc $\bar{P} \in \mathbb{Q}[T_1, \dots, T_m]$. Par ailleurs nous avons que

$$\deg(\bar{P}) = \delta \deg(P) \tag{5.151}$$

parce que $\bar{P}$ est le produit de δ «copies» de P. De plus $P = P^{\sigma_1}$ divise $\bar{P}$ donc on a bien que si $P(z) = 0$ alors $\bar{P}(z) = 0$. Le polynôme $\bar{P}$ est celui que nous cherchions. □

5.5.3 Relations coefficients racines

Théorème 5.112 (Relations coeffitients-racines)**.**
Soit le polynôme $P = a_n X^n + \dots + a_1 X + a_0$ et r_i ses n racines. Alors nous avons pour chaque $1 \leqslant k \leqslant n$ la relation

$$\sigma_k(r_1, \dots, r_n) = (-1)^k \frac{a_{n-k}}{a_n} \tag{5.152}$$

où σ_k est le k^epolynôme symétrique définit en 5.108.

Exemple 5.113
Soit le polynôme

$$P(x) = x^3 + 2x^2 + 3x + 4 \tag{5.153}$$

et ses racines que nous nommons a, b, c. Nous voudrions calculer $a^2 + b^2 + c^2$. D'abord nous décomposons $Q(a, b, c) = a^2 + b^2 + c^2$ en polynômes symétriques élémentaires : $Q(a, b, c) = \sigma_1(a, b, c)^2 - 2\sigma_2(a, b, c)$.

Mais les relations coefficients-racines (théorème 5.112) nous donnent $\sigma_1(a, b, c) = -2$ et $\sigma_2(a, b, c) = 3$, donc

$$a^2 + b^2 + c^2 = (-2)^2 - 2 \cdot 3 = -2. \tag{5.154}$$

Cela nous assure déjà qu'au moins une des solutions n'est pas réelle.

Nous pouvons en avoir une vérification directe en calculant explicitement les racines (ce qui est possible pour le degré 3) :

15. Il me semble qu'il manque la somme sur i dans [34].

```
sage: P(x)=x**3+2*x**2+3*x+4
sage: S=solve(  P(x)==0,x  )
sage: sols=[ s.rhs() for s in S  ]
sage: Q=[  s**2 for s in sols ]
sage: s=sum(Q)
sage: s.simplify_full()
-2
```

tex/frido/VAYVmNRpolynomeSym.py

Notez qu'il faut un peu chipoter pour isoler les solutions depuis la réponse de la fonction `solve`. △

En suivant le même cheminement que dans l'exemple, si P est un polynôme de degré n et si r_i sont ses racines, il est facile de calculer $Q(r_1,\dots,r_n)$ pour n'importe quel polynôme symétrique Q

Proposition 5.114 (Annulation de fonctions polynômiales[46])**.**
Soit $\mathbb{K}$ un corps et P un polynôme à n indéterminées. Nous supposons que P s'annule sur un ensemble de la forme $A_1\times\cdots\times A_n$ avec $\text{Card}(A_j)>\deg_{X_j}(P)$ pour tout j. Alors $P=0$.

De plus si $P=0$ alors tous ses coefficients sont nuls[16]*.*

Démonstration. Nous prouvons le résultat par récurrence sur le nombre n d'indéterminées. Si $n=1$, cela est le théorème 5.68. Nous classons les monômes du polynôme P par ordre de puissance de X_n et nous le factorisons :

$$P=\sum_{i=1}^{m}P_iX_n^i \tag{5.155}$$

avec $P_i\in\mathbb{K}[X_1,\dots,X_{n-1}]$. Soit $(a_1,\dots,a_{n-1})\in A_1\times\ldots\times A_{n-1}$ et posons

$$Q(T)=P(a_1,\dots,a_{n-1},T)=\sum_{i=1}^{m}P_i(a_1,\dots,a_{n-1})T^i. \tag{5.156}$$

Le polynôme Q s'annule sur A_n avec $\deg(Q)=\deg_{X_n}(P)<\text{Card}(A_n)$ et le théorème 5.68 nous donne $Q=0$. Or les coefficients des différentes puissances de T dans $Q(T)$ sont les $P_i(a_1,\dots,a_{n-1})$; ils sont donc nuls.

Nous avons montré que le polynôme P_i s'annule pour tout élément de $A_1\times\ldots\times A_{n-1}$, mais nous avons

$$\deg_{X_j}(P_i)\leqslant\deg_{X_j}P<\text{Card}(A_j), \tag{5.157}$$

donc l'hypothèse de récurrence donne $P_i=0$. Par suite, $P=0$ également. □

Exemple 5.115
Un polynôme à plusieurs variables peut s'annuler en une infinité de points sans être nul. Par exemple le polynôme $X^2+Y^2-1\in\mathbb{R}[X,Y]$ s'annule sur tout un cercle de $\mathbb{R}^2$ mais n'est pas nul, loin s'en faut. △

5.6 Polynômes cyclotomiques

5.6.1 Définitions et propriétés

Définition 5.116.
*Le **polynôme cyclotomique** d'indice n est le polynôme*

$$\phi_n(X)=\prod_{z\in\Delta_n}(X-z) \tag{5.158}$$

16. L'intérêt de cela est qu'un polynôme de $\mathbb{Z}[X_1,\dots,X_n]$ peut s'évaluer sur un élément de n'importe quel corps ; il restera le polynôme nul.

où Δ_n est l'ensemble des racines primitives de l'unité de la définition 3.103 :

$$\Delta_n = \{e^{2ik\pi/n} \text{ tel que } 0 \leqslant k \leqslant n-1 \text{ tel que } \operatorname{pgcd}(k,n) = 1\}, \tag{5.159}$$

Le polynôme ϕ_n est un polynôme unitaire de degré $\varphi(n)$ où φ est l'indicatrice d'Euler [17] $\varphi(n)$. Nous avons par exemple

$$\Delta_1 = \{1\} \tag{5.160a}$$
$$\Delta_2 = \{-1\} \tag{5.160b}$$
$$\Delta_3 = \{e^{2\pi i/3}, e^{4\pi i/3}\} \tag{5.160c}$$

et les premiers polynômes cyclotomiques sont donnés par

$$\phi_1(X) = X - 1 \tag{5.161a}$$
$$\phi_2(X) = X + 1 \tag{5.161b}$$
$$\phi_3(X) = X^2 + X + 1. \tag{5.161c}$$

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 5.117.
Le polynôme $X^n - 1$ se factorise des diverses manières suivantes :

$$X^n - 1 = \prod_{z \in \mathrm{U}_n} (X - z) = \prod_{d \mid n} \prod_{z \in \Delta_d} (X - z) = \prod_{d \mid n} \phi_d(X) \tag{5.162}$$

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 4.123 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 $\mathrm{U}_n = \bigcup_{d \mid n} \Delta_d$ du lemme 3.105 et la définition (5.158) des polynômes cyclotomiques. □

Remarque 5.118.
Notons juste pour le plaisir que dans le produit $\prod_{d \mid n} \prod_{z \in \Delta_d}$, il y a bien n termes parce que $\operatorname{Card}(\Delta_d) = \varphi(d)$ et $\sum_{d \mid n} \varphi(d) = n$ (définition 3.103 et lemme 3.105).

Proposition 5.119.
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

$$X^{n+1} - 1 = \prod_{d \mid n+1} \phi_d(X) = \phi_{n+1}(X) \cdot \underbrace{\prod_{\substack{d \mid n+1 \\ d \leqslant n}} \phi_d(X)}_{\in \mathbb{Z}[X] \text{ par récurrence}} \tag{5.163}$$

Le lemme 5.42 conclut que $\phi_{n+1} \in \mathbb{Z}[X]$. Nous avons vu $\mathbb{Z}$ comme sous anneau du corps $\mathbb{C}$. □

Proposition 5.120.
Soient $1 \leqslant m \leqslant n$ deux entiers et

$$T(X) = \frac{X^n - 1}{X^m - 1} \in \mathbb{Z}(X). \tag{5.164}$$

Alors :

17. Définie par l'équation 3.175.

(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 $\mathbb{Z}[X]$.

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

$$Q = \{\text{diviseurs de } n \text{ ne divisant pas } m\}. \tag{5.165}$$

Nous avons alors

$$X^n - 1 = \prod_{d \mid n} \phi_d(X) = \prod_{d \mid m} \phi_d(X) \cdot \prod_{q \in Q} \phi_q(X) = (X^m - 1) \cdot \prod_{q \in Q} \phi_q(X). \tag{5.166}$$

Nous avons donc

$$T(X) = \frac{X^n - 1}{X^m - 1} = \prod_{q \in Q} \phi_q(X) \in \mathbb{Z}[X]. \tag{5.167}$$

(2) Nous venons de montrer que

$$T = \prod_{q \in Q} \phi_q \in \mathbb{Z}[X]. \tag{5.168}$$

Étant donné que $m < n$ nous avons $n \in Q$ et donc

$$T = \phi_n \cdot \prod_{q \in Q \setminus \{n\}} \phi_q. \tag{5.169}$$

Par conséquent ϕ_n divise T dans $\mathbb{Z}[X]$.

□

Corollaire 5.121.
Si p est premier alors le polynôme cyclotomique ϕ_p a une bonne tête :

$$\phi_p(X) = 1 + X + \cdots + X^{p-1}. \tag{5.170}$$

Démonstration. Nous utilisons la formule du lemme 5.117 en remarquant que seuls p et 1 divisent p :

$$X^p - 1 = \prod_{d \mid p} \phi_d(X) = \phi_1(X)\phi_p(X) = (X - 1)\phi_p(X). \tag{5.171}$$

Nous pouvons simplifier par $X - 1$ en utilisant la formule du lemme 4.125(2) :

$$1 + X + \cdots + X^{p-1} = \phi_p(X) \tag{5.172}$$

□

Proposition 5.122 (Irréductibilité des polynômes cyclotomiques[47])**.**
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 5.63 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 5.45 nous montre que $h \in \mathbb{Z}[X]$ parce que ϕ_n, f et g ont des coefficients entiers. Nous avons

$$f(\xi) = g(\xi^l) = 0. \tag{5.173}$$

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

$$\psi = fk \tag{5.174}$$

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

$$\bar{\psi}(X) = \bar{g}(X^l) = \bar{g}(X)^l. \tag{5.175}$$

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 5.181.

Corollaire 5.123.
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 5.122, il est unitaire par définition et contient le monôme $X - e^{2i\pi/q}$, donc il est annulateur. Annulateur, irréductible et unitaire, le corollaire 5.67 en fait le polynôme minimal de ω. □

Théorème 5.124.
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

$$P = \prod_{i=1}^{d}(X - \xi_i) \tag{5.176}$$

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

$$g_q(X) = \prod_{i=1}^{d}\big(X - (\xi_i)^q\big), \tag{5.177}$$

et en particulier $g_1 = P$, et nous développons

$$g_q(X) = X^n + C_{1,q}X^{n-1} + \cdots + C_{n,q} \tag{5.178}$$

où

$$C_{k,q} = (-1)^k \sum_{1\leqslant i_1<\ldots<i_k\leqslant d} (\xi_{i_1}\ldots\xi_{i_k})^q. \tag{5.179}$$

Nous introduisons aussi les polynômes

$$F_{k,q}(X_1,\ldots,X_n) = (-1)^k \sum_{1\leqslant i_1<\ldots<i_k\leqslant d} (X_{i_1}\ldots X_{i_k})^q \tag{5.180}$$

qui sont des polynômes symétriques. Ils vérifient deux propriétés. La première est que

$$C_{r,q} = F_{r,q}(\xi_1,\ldots,\xi_n), \tag{5.181}$$

et la seconde est que les polynômes $F_{r,1}$ sont les polynômes symétriques élémentaires à un coefficients près. Le théorème 5.109 nous donne alors des polynômes $G_{k,q}\in\mathbb{Z}[X_1,\ldots,X_n]$ tels que

$$F_{k,q}(X_1,\ldots,X_n) = G_{k,q}\big(F_{1,1}(X_1,\ldots,X_n),\ldots,F_{k,1}(X_1,\ldots,X_n)\big). \tag{5.182}$$

Nous savons que

$$|C_{k,q}| \leqslant \sum_{1\leqslant i_1<\ldots<i_k<d} 1 = \binom{d}{k}. \tag{5.183}$$

Donc g_q fait partie de l'ensemble fini des polynômes dans $\mathbb{Z}[q]$ dont tous les coefficients sont bornée en valeur absolue par

$$\max_{k=1,\ldots,d}\binom{d}{k}. \tag{5.184}$$

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

$$\{\xi_k^q \text{ tel que } q\in\mathbb{N}\}. \tag{5.185}$$

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

$$\xi_k^N = 1 \tag{5.186}$$

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

$$X^N - 1 = \prod_{d\,|\,N}\phi_d(X), \tag{5.187}$$

les polynômes cyclotomiques sont les seuls facteurs irréductibles de X^N-1. Donc P est un polynôme cyclotomique. □

5.6.2 Nombres premiers

Lemme 5.125 ([48])**.**
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

$$B(X) = \prod_{\substack{d \mid n \\ d \neq n}} \phi_d(X), \tag{5.188}$$

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 5.158, les polynômes cyclotomiques sont scindés (dans $\mathbb{C}$), donc en particulier les polynômes ϕ_n et B sont scindés et dons premiers entre eux, dans $\mathbb{C}$ et a fortiori dans $\mathbb{Q}$. Par Bézout (corollaire 4.77), il existe $U, V \in \mathbb{Q}[X]$ tels que

$$U\phi_n + VB = 1. \tag{5.189}$$

Si nous prenons $a \in \mathbb{Z}$ tel que $U' = aU$ et $V' = aV$ soient tous deux dans $\mathbb{Z}[X]$, alors nous avons

$$U'\phi_n + V'B = a, \tag{5.190}$$

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 construis 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 (5.190) en a :

$$U'(a)\phi_n(a) + V'(a)B(a) = a. \tag{5.191}$$

Le nombre p ne divisant pas a, mais divisant $\phi_n(a)$, il ne peux pas diviser $B(a)$ [18]. 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

$$X^d - 1 = \prod_{d' \mid d} \phi_{d'}, \tag{5.192}$$

et cela est une partie du produit

$$\prod_{\substack{d \mid n \\ d \neq n}} \phi_d. \tag{5.193}$$

Vu que p ne divise aucun des $\phi_d(a)$ de ce dernier produit, a fortiori, il ne divise pas le produit 5.192, et donc pas $a^d - 1$. □

Lemme 5.126.
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. Soit $n \geqslant 1$ et les nombres p, a donnés par le lemme 5.125. 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

$$a^d - 1 = \prod_{d' \mid d} \phi_{d'}(a). \tag{5.194}$$

18. C'est pour pouvoir dire ça que l'on a choisit $V' \in \mathbb{Z}[X]$ de telle sorte que $V'(a)$ soit dans $\mathbb{Z}$

Par construction de a et p, nous avons

$$[\phi_{d'}(a)]_p \neq 0 \tag{5.195}$$

Vu que $\mathbb{Z}/p\mathbb{Z}$ est intègre, le produit est également non nul, c'est à dire

$$\Big[\prod_{d'\,|\,d} \phi_{d'}(a)\Big]_p \neq 0, \tag{5.196}$$

et donc $[a]_p^a \neq 1$. Nous avons donc montré que si $d \neq n$ divise n, alors nous avons en même temps

$$[a]_p^n = 1 \tag{5.197}$$

et

$$[a]_p^d \neq 1. \tag{5.198}$$

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 5.127 (Forme faible du théorème de Dirichlet [34])**.**
Pour tout $n \geqslant 1$, il existe une infinité de nombres premiers dans $[1]_n$.

Démonstration. Le lemme 5.126 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

$$N = np_1 \dots p_r. \tag{5.199}$$

Nous utilisons maintenant le lemme 5.126 avec ce N, c'est à dire qu'on a un nombre premier de la forme

$$p = 1 + kN = 1 + knp_1 \dots p_r. \tag{5.200}$$

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 $p_1, \dots, p_r$. □

5.6.3 Théorème de Wedderburn

Théorème 5.128 (Théorème de Wedderburn[41])**.**
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 5.46 nous avons

$$\text{Card}(\mathbb{K}) = q^n \tag{5.201}$$

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

$$Z_x = \{a \in \mathbb{K} \text{ tel que } ax = xa\}. \tag{5.202}$$

Le centre Z est un sous corps de Z_x, donc il existe $d(x)$ tel que

$$\text{Card}(Z_x) = q^{d(x)}. \tag{5.203}$$

De la même manière, Z_x est un sous corps de $\mathbb{K}$, donc il existe $m(x)$ tel que

$$\text{Card}(\mathbb{K}) = \text{Card}(Z_x)^{m(x)}. \tag{5.204}$$

En mettant bout à bout nous avons

$$q^n = \text{Card}(Z_x)^{m(x)} = q^{d(x)m(x)}, \tag{5.205}$$

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 :

$$\varphi(k)x = kxk^{-1}. \tag{5.206}$$

Nous notons $\mathcal{O}_x$ l'orbite de $x \in \mathbb{K}^*$ pour cette action, et $\text{Fix}(x)$ son stabilisateur. Nous avons

$$Z_y = \text{Fix}(y) \cup \{0\} \tag{5.207}$$

parce que Z_y et $\text{Fix}(y)$ ont les mêmes définitions, sauf que $\text{Fix}(y)$ est dans $\mathbb{K}^*$ alors que Z_y est dans $\mathbb{K}$. Nous avons donc

$$\text{Card}\big(\text{Fix}(y)\big) = \text{Card}(Z_y) - 1 = q^{d(y)} - 1. \tag{5.208}$$

Nous avons $\text{Card}(\mathcal{O}_x) = 1$ si et seulement si $\mathcal{O}_x = \{x\}$ si et seulement si $\text{Fix}(x) = \mathbb{K}^*$ 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

$$\{z_1\}, \dots, \{z_{q-1}\} \tag{5.209}$$

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 (3.148) :

$$\text{Card}(\mathbb{K}^*) = \text{Card}(Z^*) + \sum_{i=1}^{r} \frac{\text{Card}(\mathbb{K}^*)}{\text{Card}(\text{Fix}(y_i))}, \tag{5.210}$$

mais $\text{Card}(Z^*) = q - 1$, $\text{Card}(\mathbb{K}^*) = q^n - 1$ et $\text{Card}\big(\text{Fix}(y_i)\big) = q^{d(y_i)} - 1$, donc

$$q^n - 1 = (q - 1) + \sum_{i=1}^{r} \frac{q^n - 1}{q^{d(y_i)} - 1}. \tag{5.211}$$

Nous considérons la fraction rationnelle

$$F(X) = (X^n - 1) - \sum_{i=1}^{r} \frac{X^n - 1}{X^{d(y_i)} - 1}. \tag{5.212}$$

Étant donné que $d(y_i)$ divise n, nous avons, contrairement aux apparences, que $F \in \mathbb{Z}[X]$ par la proposition 5.120(1).

Nous pouvons exploiter un peu mieux la proposition 5.120 en remarquant que $d(y_i) < n$ parce que sinon $\text{Card}(Z_{y_i}) = \text{Card}(\mathbb{K})$, ce qui signifierait que $y_i \in Z$, ce qui nous avions exclu. Par conséquent le polynôme cyclotomique ϕ_n divise

$$\frac{X^n - 1}{X^{d(y_i)} - 1} \tag{5.213}$$

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 :

$$F(q) = Q(q)\phi_n(q) = q - 1. \tag{5.214}$$

En effet nous avons $F(q) = q - 1$ par construction : comparer (5.211) avec (5.212). Évidemment $q \neq 1$ parce que si $q = 1$ alors $\text{Card}(\mathbb{K}) = 1$ et le théorème est trivial. Par ailleurs $Q(q)$ est un entier (parce que $Q \in \mathbb{Z}[X]$ et $q \in \mathbb{N}$) et $Q(q) \neq 0$, parce qu'à droite de (5.214) nous avons $q - 1 \neq 0$. Nous avons donc $|Q(q)| \geqslant 1$ et donc

$$|\phi_n(q)| \leqslant q - 1. \tag{5.215}$$

Par définition du polynôme cyclotomique nous avons

$$|\phi_n(q)| = \prod_{z\in\Delta_n} |q-z|. \tag{5.216}$$

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 5.1, la distance entre z_0 et q doit être strictement plus grande que $q-1$ parce que $q-1$ est le minimum de la distance entre le cercle trigonométrique et q, et n'est atteint qu'en $z=1$.

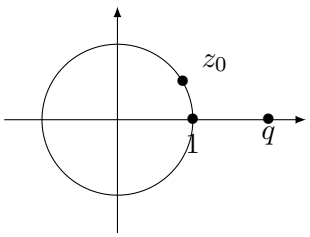


FIGURE 5.1 – Nous devons avoir $|z_0 - q| > q-1$.

Nous avons ainsi obtenu une contradiction, et nous concluons que le corps $\mathbb{K}$ est commutatif. □

5.7 Corps finis

Si vous cherchez des choses à propos de RSA, c'est à la section 10.4.

5.7.1 Existence, unicité

Nous avons déjà défini le corps fini $\mathbb{F}_p$ lorsque p est un nombre premier dans la section 5.1.3. Le théorème suivant sert à définir $\mathbb{F}_{p^n}$ lorsque p est premier.

Théorème 5.129.
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 3.55 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

$$X^{q-1} - 1 \tag{5.217}$$

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 5.78.

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

$$(\alpha\beta)^q = \alpha^q\beta^q = \alpha\beta \tag{5.218}$$

parce que $\alpha^q = \alpha$. Le produit $\alpha\beta$ est donc encore dans $\mathbb{E}$. Pour la somme,

$$(\alpha+\beta)^q = (\alpha+\beta)^{p^n} = \Big((\alpha+\beta)^p\Big)^{p^{n-1}} = (\alpha^p+\beta^p)^{p^{n-1}} = \ldots = \alpha^{p^n} + \beta^{p^n} = \alpha+\beta. \tag{5.219}$$

En ce qui concerne l'inverse,

$$(\alpha^{-1})^q = (\alpha^q)^{-1} = \alpha^{-1}. \tag{5.220}$$

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

$$P'(X) = qX^{q-1} - 1 = -1 \tag{5.221}$$

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 5.129 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 [49, 50, 51] et de wikipedia

Proposition 5.130 ([51])**.**
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 3.55. Soit d un diviseur de $q-1$ et

$$H_d^* = \{x \text{ d'ordre } d \text{ dans } \mathbb{K}^*\} \tag{5.222a}$$
$$H_d = \{\text{racines de } X^d - 1 \text{ dans } \mathbb{K}\}. \tag{5.222b}$$

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

$$\begin{aligned} \phi\colon \mathbb{Z}/d\mathbb{Z} &\to H_d \\ n &\mapsto a^n \end{aligned} \tag{5.223}$$

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 3.107, et H_d étant isomorphe à $\mathbb{Z}/d\mathbb{Z}$ nous avons

$$\text{Card}(H_d^*) = \varphi(d). \tag{5.224}$$

Par conséquent si H_d^* n'est pas vide, son cardinal est $\varphi(d)$. Nous avons

$$q - 1 = \text{Card}(\mathbb{K}^*) \tag{5.225a}$$
$$= \text{Card}\Big(\bigcup_{d\,|\,q-1} H_d^*\Big) \tag{5.225b}$$
$$= \sum_{d\,|\,q-1} \text{Card}(H_d^*) \tag{5.225c}$$
$$\leqslant \sum_{d\,|\,q-1} \varphi(d) \tag{5.225d}$$
$$= q - 1 \tag{5.225e}$$

où nous avons utilisé le lemme 3.105. 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 5.131.
Si p est un nombre premier, alors

$$(\mathbb{Z}/p\mathbb{Z})^* \simeq \mathbb{Z}/(p-1)\mathbb{Z}. \tag{5.226}$$

L'isomorphisme est un isomorphisme de groupe (abéliens). À gauche multiplicatif et à droite additif.

Démonstration. La proposition 5.130 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 5.132.
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 3.55 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

$$\prod_{a\in\mathbb{K}}(X-a) \tag{5.227}$$

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 5.133
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 fait pas regarder le théorème 5.68 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 5.134
Si $\mathbb{F}_3 = \mathbb{Z}/3\mathbb{Z}$ est le[19] corps à 3 éléments, alors le polynôme $P(X) = X^3 - X$ s'évalue à zéro pour tout $x \in \mathbb{F}_3$ (proposition 5.132.) mais il n'est pas le polynôme nul. △

5.7.2 Symboles de Legendre et carrés

Source : [52].
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 5.135.
*Soit $n \in \mathbb{N}$ et $p > 2$ un nombre premier. Le **symbole de Legendre** par*

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & \text{si } p \text{ divise } n \\ 1 & \text{si } n \text{ est un carré dans } \mathbb{F}_p \\ -1 & \text{sinon.} \end{cases} \tag{5.228}$$

19. Le singulier est justifié par le théorème 5.129, mais ça n'a pas d'importance ici.

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 5.136.
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

$$\left(\frac{n}{p}\right) = n^{(p-1)/2} \mod p. \tag{5.229}$$

Démonstration. Nous considérons l'application

$$\begin{aligned} \psi\colon \mathbb{F}_p^* &\to \mathbb{F}_p^* \\ x &\mapsto x^2. \end{aligned} \tag{5.230}$$

C'est un morphisme de groupes multiplicatifs et $\ker\psi = \{-1, 1\}$. Étant donné que $p > 2$, nous avons alors

$$\text{Card}(\ker\psi) = 2 \tag{5.231}$$

parce que $1 \neq -1$. Évidemment l'ensemble des carrés dans $\mathbb{F}_p^*$ est l'image de ψ. Le premier théorème d'isomorphisme 3.22(3) nous permet alors de conclure que

$$\text{Card}(\text{Image}(\psi)) = \frac{\text{Card}(\mathbb{F}_p^*)}{2}. \tag{5.232}$$

Ceci prouve la première assertion.

Par le petit théorème de Fermat (théorème 5.13), 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

$$X^{(p-1)/2} = \pm 1. \tag{5.233}$$

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 (5.229). 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

$$n^{(p-1)/2} = n^{p-1} = 1 = \left(\frac{n}{p}\right). \tag{5.234}$$

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

$$n^{(p-1)/2} = -1 = \left(\frac{n}{p}\right). \tag{5.235}$$

□

Corollaire 5.137.
Si $a, b \in \mathbb{N}$ et si $p > 2$ est un nombre premier, alors

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{5.236}$$

Démonstration. Par la formule (5.229),

$$\left(\frac{ab}{p}\right) = (ab)^{(p-1)/2} = a^{(p-1)/2}b^{(p-1)/2} = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right). \tag{5.237}$$

□

Soit un nombre premier $q > 2$ et $\mathbb{A}$, un anneau de caractéristique p. Si $\alpha \in \mathbb{A}$ vérifie

$$1 + \alpha + \cdots + \alpha^{q-1} = 0, \tag{5.238}$$

nous définissons la **somme de Gauss** par

$$\tau = \sum_{x \in \mathbb{F}_q} \left(\frac{i}{q}\right) \alpha^i = \sum_{x=1}^{q-1} \left(\frac{x}{q}\right) \alpha^i. \tag{5.239}$$

Notons que la somme de Gauss dépend de q et du α choisis.

Proposition 5.138.
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*

$$\tau^p = \left(\frac{p}{q}\right) \tau. \tag{5.240}$$

(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

$$\alpha^q - 1 = (\alpha - 1)(1 + \alpha + \cdots + \alpha^{q-1}) = 0 \tag{5.241}$$

par définition de α. Nous calculons

$$\epsilon(q)\tau^2 = \epsilon(q) \sum_{x,y \in \mathbb{F}_q} \left(\frac{x}{q}\right)\left(\frac{y}{q}\right) \alpha^{x+y} \tag{5.242a}$$

$$= \sum_{x,y \in \mathbb{F}_q} \left(\frac{-xy}{q}\right) \alpha^{x+y}. \tag{5.242b}$$

$$= \sum_{z \in \mathbb{F}_q} \sum_{y \in \mathbb{F}_q} \left(\frac{-(z-y)y}{q}\right) \alpha^z \tag{5.242c}$$

$$= \sum_{z \in \mathbb{F}_q} s_z \alpha^z \tag{5.242d}$$

Justifications :

— Pour obtenir (5.242b) nous avons utilisé le corollaire 5.137.
— (5.242c) est un changement de variable $z = x + y$ dans la somme sur x.
— Pour (5.242d) nous avons posé

$$s_z = \sum_{y \in \mathbb{F}_q} \left(\frac{-(z-y)y}{q}\right). \tag{5.243}$$

Nous avons

$$s_0 = \sum_{y \in \mathbb{F}_q} \left(\frac{y^2}{q}\right). \tag{5.244}$$

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

$$\begin{aligned} \mathbb{F}_q^* &\to \mathbb{F}_q \backslash \{1\} \\ k &\mapsto 1 - zy^{-1} \end{aligned} \tag{5.245}$$

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

$$\sum_{y\in\mathbb{F}_q}\left(\frac{y(z-y)}{q}\right) = \sum_{y\in\mathbb{F}_q}\left(\frac{y^2(y^{-1}z-1)}{q}\right) \tag{5.246a}$$

$$= \sum_{y\in\mathbb{F}_q}\left(\frac{y^{-1}z-1}{q}\right) \tag{5.246b}$$

$$= \sum_{t\in\mathbb{F}_q\setminus\{1\}}\left(\frac{t}{q}\right) \tag{5.246c}$$

$$= \underbrace{\sum_{t\in\mathbb{F}_q}\left(\frac{y}{q}\right)}_{=0} - \left(\frac{1}{1}\right) \tag{5.246d}$$

$$= -1 \tag{5.246e}$$

parce qu'il y a autant de carrés que de non carrés dans $\mathbb{F}_q^*$ (proposition 5.136). En résumé nous avons

$$\epsilon(q)\tau^2 = \sum_{z\in\mathbb{F}_q} s_z\alpha^z \tag{5.247}$$

où

$$s_z = \begin{cases} q-1 & \text{si } z = 0 \\ -1 & \text{sinon.} \end{cases} \tag{5.248}$$

Cela donne

$$\epsilon(q)\tau^2 = (q-1) - \underbrace{(\alpha + \cdots + \alpha^{q-1})}_{=-1} = q \tag{5.249}$$

où nous avons utilisé l'hypothèse sur α. Donc $\epsilon(q)\tau^2 = q$, et étant donné que $\epsilon(q) = \pm 1$ nous concluons

$$\tau^2 = \epsilon(q)q. \tag{5.250}$$

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,

$$\tau^p = \left(\sum_{x\in\mathbb{F}_q}\left(\frac{x}{q}\right)\alpha^x\right)^p = \sum_{x\in\mathbb{F}_q}\left(\frac{x}{q}\right)^p\alpha^{px}. \tag{5.251}$$

Étant donné que $\left(\frac{x}{q}\right) = \pm 1$ et que p est impair, nous avons

$$\left(\frac{x}{q}\right)^p = \left(\frac{x}{q}\right). \tag{5.252}$$

Du coup nous avons

$$\left(\frac{p}{q}\right)\tau^p = \sum_{x\in\mathbb{F}_p}\left(\frac{xp}{q}\right)\alpha^{px}. \tag{5.253}$$

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 :

$$\left(\frac{p}{q}\right)\tau^p = \sum_{x\in\mathbb{F}_p}\left(\frac{x}{q}\right)\alpha^x = \tau. \tag{5.254}$$

Nous trouvons alors que

$$\tau^p = \left(\frac{p}{q}\right)\tau. \tag{5.255}$$

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 3.32) nous donne a et b tels que

$$ap + ba = 1. \tag{5.256}$$

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 5.139 (Loi de réciprocité quadratique)**.**
Soient deux nombres premiers distincts $p, q \geqslant 3$. Alors

$$\left(\frac{p}{q}\right) = (-1)^{\frac{(p-1)(q-1)}{4}} \left(\frac{q}{p}\right). \tag{5.257}$$

Démonstration. Soit ϕ_q le polynôme $A + X + \cdots + X^{q-1}$ et l'anneau

$$\mathbb{A} = \mathbb{F}_p[X]/(\phi_q). \tag{5.258}$$

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

$$\tau = \sum_{i \in \mathbb{F}_q} \left(\frac{i}{q}\right) \alpha^i. \tag{5.259}$$

Notons que cela est un élément de $\mathbb{A}$ et plus précisément une (classe de) polynôme de degré zéro dans $\mathbb{A}$. Donc les coefficients de α doivent être compris comme des éléments de $\mathbb{F}_p$. Nous savons (proposition 5.138) que

$$\tau^2 = \left(\frac{-1}{q}\right), \tag{5.260}$$

et en utilisant la formule (5.229) nous trouvons

$$\left(\frac{\tau^2}{p}\right) = (\tau^2)^{(p-1)/2} \mod p = \tau^{p-1} \mod p \tag{5.261}$$

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 (5.240) nous avons

$$\underbrace{\left(\frac{\tau^2}{p}\right)}_{\tau^{p-1}} \tau = \tau^p = \left(\frac{p}{q}\right) \tau \tag{5.262}$$

Vu que τ est inversible, nous écrivons

$$\left(\frac{\tau^2}{p}\right) = \left(\frac{p}{q}\right) \tag{5.263}$$

Nous utilisons maintenant la formule (5.229) sur le membre de gauche avec $n = \tau^2 = \left(\frac{-1}{q}\right)$:

$$\left(\frac{p}{q}\right) = \left(\frac{-1}{q}\right)^{\frac{q-1}{2}} \left(\frac{q}{p}\right). \tag{5.264}$$

Toujours avec la même formule nous pouvons substituer $\left(\frac{-1}{q}\right)$ par $(-1)^{(q-1)/2}$ et obtenir

$$\left(\frac{p}{q}\right) = (-1)^{\frac{(q-1)}{2}\frac{(p-1)}{2}}. \tag{5.265}$$

□

Lemme 5.140.
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 5.136. Pour l'unicité, soit $\alpha\colon \mathbb{F}_p^* \to \{-1,1\}$ un morphisme surjectif (c'est à dire non trivial). Étant donné que

$$\mathbb{F}_p^* = \ker(\alpha) \cup -\ker(\alpha), \tag{5.266}$$

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 5.130), alors $a^2 \in S$ par le lemme 3.60. 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 5.136 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,

$$\ker(\alpha) = (\mathbb{F}_p^*)^2. \tag{5.267}$$

Au final, pour $y \in \mathbb{F}_p^*$,

$$\alpha(y) = \begin{cases} 1 & \text{si } y \text{ est un carré} \\ -1 & \text{sinon.} \end{cases} \tag{5.268}$$

Cela est bien la définition des symboles de Legendre. □

Proposition 5.141.
Pour p premier nous avons

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \textit{si } p \in [1]_8 \textit{ ou } p \in [7]_8 \\ -1 & \textit{sinon.} \end{cases} \tag{5.269}$$

Démonstration. Soit le polynôme

$$X^4 + 1 \in \mathbb{F}_p[X] \tag{5.270}$$

et α, une racine dans une fermeture de $\mathbb{F}_p$. Nous posons $\theta = \alpha + \alpha^{-1}$ et nous calculons

$$\theta^2 = (\alpha + \alpha^{-1})(\alpha + \alpha^{-1}) = \alpha^2 + 2 + (\alpha^2)^{-1} = \alpha^2 + 2 - \alpha^2 = 2 \tag{5.271}$$

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 5.61, 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 :

$$1, \alpha, \alpha^2, \alpha^3, -1, -\alpha, -\alpha^2, -\alpha^3, 1, \alpha, \dots \tag{5.272}$$

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

$$\alpha^2 = \alpha \tag{5.273a}$$
$$\alpha^3 = \alpha \tag{5.273b}$$
$$-\alpha = \alpha \tag{5.273c}$$
$$-\alpha^3 = \alpha \tag{5.273d}$$

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 5.61, 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,

$$\theta^p = (\alpha + \alpha^{-1})^p = \alpha^p + \alpha^{-p}. \tag{5.274}$$

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

$$\theta^p = \alpha^{1+8k} + (\alpha^{-1})^{1+8k} = \alpha + \alpha^{-1} = \theta, \tag{5.275}$$

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

$$\alpha^6 + 1 = \alpha^4 + \alpha^2, \tag{5.276}$$

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. □

5.7.3 Théorème de Chevalley-Warning

Lemme 5.142.
Soit $\mathbb{K}$ un corps de caractéristique p et de cardinal q. Pour $m \in \mathbb{N}$ nous définissons

$$S_m = \sum_{x \in \mathbb{K}} x^m. \tag{5.277}$$

Alors nous avons

$$S_m \mod p = \begin{cases} -1 & \text{si } m \geqslant 1 \text{ et } m \text{ divisible par } q-1 \\ 0 & \text{sinon.} \end{cases} \tag{5.278}$$

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 4.21).

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

$$x^m = x^{k(q-1)} = 1 \tag{5.279}$$

parce que $\mathbb{K}^*$ est cyclique et $x^{q-1} = 1$ par le petit théorème de Fermat (théorème 5.13). Par conséquent nous avons

$$\sum_{x \in \mathbb{K}} x^m = \sum_{x \in \mathbb{K}^*} 1 = q - 1. \tag{5.280}$$

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

$$\begin{aligned} \varphi\colon \mathbb{K}^* &\to \mathbb{K}^* \\ x &\mapsto yx \end{aligned} \tag{5.281}$$

est une bijection [20]. 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

$$z = \varphi(a^{l-k}). \tag{5.282}$$

20. Notons que nous n'avons pas réellement besoin que y soit un générateur. Nous n'utilisons seulement le fait que $y^m \neq 1$ et $y \neq 0$.

Nous pouvons maintenant faire le calcul.

$$S_m = \sum_{x\in\mathbb{K}^*} x^n = \sum_{x\in\mathbb{K}^*} (yx)^m = y^m \sum_{x\in\mathbb{K}^*} x^m = y^m S_m. \tag{5.283}$$

Étant donné que $y^m \neq 1$, la seule solution est $S_m = 0$. □

Théorème 5.143 (Chevalley-Warning)**.**
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 :

$$V = \{x \in \mathbb{K}^n \text{ tel que } P_1(x) = \dots = P_r(x) = 0\}. \tag{5.284}$$

Alors $\text{Card}(V) = 0 \mod p$.

Démonstration. Nous considérons le polynôme

$$P = \prod_{i=1}^{r} (1 - P_i^{q-1}). \tag{5.285}$$

Montrons que

$$P(x) = \begin{cases} 1 & \text{si } x \in V \\ 0 & \text{sinon.} \end{cases} \tag{5.286}$$

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

$$\deg(P) = \sum_{i=1}^{r} (q-1)\deg(P_i) < n(q-1). \tag{5.287}$$

Pour un polynôme $Q \in \mathbb{K}[X_1, \dots, X_n]$, nous définissons

$$\int Q = \sum_{x\in\mathbb{K}^n} Q(x). \tag{5.288}$$

Nous avons immédiatement

$$\int P = \sum_{x\in\mathbb{K}^n} P(x) = \sum_{x\in V} 1 = \text{Card}(V) \mod p. \tag{5.289}$$

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

$$P = \sum_m c_m X_1^{m_1} \dots X_n^{m_n} \tag{5.290}$$

où la somme s'étend sur les $m \in \mathbb{N}^n$ tels que $c_m \neq 0$. Nous avons

$$\int P = \sum_{\in\mathbb{K}^n} \sum_m c_m x_1^{m_1} \dots x_n^{m_n} \tag{5.291a}$$

$$= \sum_m c_m \left(\sum_{x\in\mathbb{K}^n} x_1^{m_1} \dots x_n^{m_n} \right) \tag{5.291b}$$

$$= \sum_m c_m S_{m_1} \dots S_{m_n}. \tag{5.291c}$$

Le terme de plus haut degré dans la décomposition (5.290) 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

$$\sum_{i=1}^{n} m_i < n(q-1). \tag{5.292}$$

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

$$\sum_{m\in\mathbb{N}^n} c_m S_{m_1} \dots S_{m_n} \tag{5.293}$$

ont un facteur nul. □

Corollaire 5.144.
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 5.143. 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 5.145
Nous considérons les polynômes

$$P_1(x, y, t, u) = xy + x + ux \tag{5.294a}$$
$$P_2(x, y, t, u) = x + y - 3t. \tag{5.294b}$$

La somme de leurs degrés est 3 et ce sont des polynômes à 4 variables. Nous devons donc avoir, en vertu du corollaire 5.144, 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 $P_1 = P_2 = 0$. △

5.7.4 Contenu d'un polynôme

Lemme 5.146 (de Gauss[2, 53])**.**
Soient $P, Q \in \mathbb{Z}[X]$. Alors

$$c(PQ) = c(P)c(Q) \tag{5.295}$$

où c est le contenu, définition 4.112.

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 :

$$p \mid a_0, p \mid a_1, \dots, p \mid a_{i_0-1}, p \nmid a_{i_0} \tag{5.296}$$

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 :

$$a_{i_0}b_{j_0} + \sum_{\substack{i+j=i_0+j_0 \\ i<i_0 \text{ ou } j<j_0}} a_i b_j. \tag{5.297}$$

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 :

$$c(P_1Q_1) = 1. \tag{5.298}$$

Étant donné que

$$P_1Q_1 = \frac{1}{c(P)c(Q)}PQ, \tag{5.299}$$

nous avons

$$c(PQ) = c(P)c(Q)c(P_1Q_1) = c(P)c(Q). \tag{5.300}$$

□

5.7.5 Théorème de l'élément primitif

Définition 5.147.
*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 5.148 (de l'élément primitif)**.**
Si $\mathbb{K}$ est un corps fini, toute extension finie de $\mathbb{K}$ est simple[21].
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 5.130 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 5.149.
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 5.150.
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 :

$$(\alpha + \beta)^p = \sum_k \binom{k}{p} \alpha^k \beta^{p-k} \tag{5.301}$$

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 $(\alpha + \beta)^{p^n}$.

21. Définition 5.57.

Lemme 5.151.
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 $i = 0,\dots,k-1\}$.

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 5.152 (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

$$\begin{aligned} \phi\colon \mathbb{F}_p[X]/(P) &\to \mathbb{K} \\ \bar{X} &\mapsto \alpha \end{aligned} \tag{5.302}$$

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[22]*.*

(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 5.130. Si α un générateur de $\mathbb{K}^*$ alors

$$\mathbb{K} = \mathbb{F}_p[\alpha]. \tag{5.303}$$

Soit ℓ le plus grand entier tel que l'ensemble

$$\{1, \alpha, \cdots, \alpha^{\ell-1}\} \subset \mathbb{K} \tag{5.304}$$

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

$$\alpha^\ell + a_{\ell-1}\alpha^{\ell-1} + \cdots + a_0 = 0. \tag{5.305}$$

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 $\mathbb{K}$,

$$\mathbb{K} = \text{Span}\{1, \alpha, \dots, \alpha^{\ell-1}\} \tag{5.306}$$

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

$$\text{Card}(\mathbb{K}) = p^n = q \tag{5.307}$$

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

$$\{\alpha^{\ell-1}, \dots, 1\} \tag{5.308}$$

est libre.

22. Définition 4.114.

Montrons que l'application

$$\begin{aligned}\phi\colon \mathbb{F}_p[X]/(P) &\to \mathbb{K}\\ \bar{X} &\mapsto \alpha\end{aligned} \tag{5.309}$$

est un isomorphisme. Pour l'injectivité, deux éléments $Q_1, Q_2 \in \mathbb{F}_p[X]/(P)$ s'écrivent

$$Q_1 = \sum_{k=0}^{n-1} a_k \bar{X}^k \tag{5.310a}$$

$$Q_2 = \sum_{k=0}^{n-1} b_k \bar{X}^k. \tag{5.310b}$$

Dans ce cas si $\phi(Q_1) = \phi(Q_2)$ alors

$$\phi(Q_1) = \sum_{k=0}^{n-1} a_k \alpha^k = \phi(Q_2) = \sum_{k=0}^{n-1} b_k \alpha^k. \tag{5.311}$$

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

$$\{\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}}\} \tag{5.312}$$

est l'ensemble des racines distinctes de P. Pour cela nous posons

$$P(X) = \sum_{k=0}^{n} a_k X^k \tag{5.313}$$

avec $a_k \in \mathbb{F}_p$. D'abord α est une racine de P. En effet

$$P(\bar{X}) = \sum_k a_k \bar{X}^k = 0 \tag{5.314}$$

parce que cette somme est calculée dans $\mathbb{F}_p[X]/(P)$. En appliquant l'isomorphisme ϕ à l'égalité (5.314) nous trouvons

$$0 = \phi\big(P(\bar{X})\big) = \sum_k a_k \phi(\bar{X}^k) = \sum_k a_k \alpha^k. \tag{5.315}$$

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 (5.312). Étant donné que pour tout x dans $\mathbb{F}_p$ nous avons $x^p = x$, nous avons aussi

$$P(X^p) = \sum_k a_k (X^p)^k = \sum_k a_k^p (X^p)^k = \sum_k (a_k X^k)^p \tag{5.316}$$

alors que nous savons que $x \mapsto x^p$ est un automorphisme de $\mathbb{F}_p$ par la proposition 4.24. Par conséquent

$$P(X^p) = \sum_k (a_k X^k)^p = \left(\sum_k a_k X^k \right)^p = P(X)^p. \tag{5.317}$$

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

$$\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}} \tag{5.318}$$

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

$$\{\alpha, \alpha^p, \alpha^{p^2}, \ldots, \alpha^{p^{n-1}}\} \tag{5.319}$$

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 $\alpha^{q-1} = e = \alpha^{p^n-1}$,

$$\beta^q = (\alpha^{p^n})^k \tag{5.320a}$$
$$= \left(\alpha^{p^n-1}\alpha\right)^k \tag{5.320b}$$
$$= \left(\alpha^{q-1}\alpha\right)^k \tag{5.320c}$$
$$= \alpha^k \tag{5.320d}$$
$$= \beta. \tag{5.320e}$$

Cela signifie que $\beta^q = \beta$ et donc que β est racine de $X^q - X$. □

Corollaire 5.153.
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 (5.302). Soit 1_q la classe du polynôme 1 modulo P, nous considérons le morphisme

$$\begin{aligned} \mu\colon \mathbb{Z} &\to \mathbb{F}_q \\ n &\mapsto n1_q. \end{aligned} \tag{5.321}$$

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 5.154.
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 5.155.
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 $P = \sum_k a_k X^k$,

$$P(\alpha^p) = \sum_k (a_k\alpha^k)^p = \big(\sum_k a_k\alpha^k\big)^p = 0 \tag{5.322}$$

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 3.55 indique que $\alpha^{p^n} = \alpha$. En particulier avec $r = p^{n-k}$ nous avons

$$\beta^r = \alpha^{rp^k} = \alpha^{p^n} = \alpha. \tag{5.323}$$

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 5.156.
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 5.157.
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 5.158.
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 5.152. L'élément a est en particulier une racine de $X^q - X$. Par ailleurs P divise $X^q - X$ par le lemme 5.156.

Nous avons aussi

$$X^q - X = \prod_{b \in \mathbb{L}} (X - b) \tag{5.324}$$

par la proposition 5.132. 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

$$\mathbb{F}_p[X]/P \simeq \mathbb{F}_p[X]/Q \simeq \mathbb{K}. \tag{5.325}$$

Nous montrons maintenant que $\mathbb{F}_p[X]/Q \simeq \mathbb{L}$ par l'application

$$\begin{aligned} \phi\colon \mathbb{F}_p[X]/Q &\to \mathbb{L} \\ \bar{X} &\mapsto b \end{aligned} \tag{5.326}$$

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. □

5.7.6 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 5.129) et que nous le notons $\mathbb{F}_q$. Le théorème 5.152 nous incite à l'écrire sous la forme

$$\mathbb{F}_q = \mathbb{F}_p[X]/(P) \tag{5.327}$$

pour un certain polynôme irréductible $P \in \mathbb{F}_p[X]$.

5.7.6.1 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 5.159.
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 5.38, $\mathbb{K}$ 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 $\bar{X}$.

□

5.7.6.2 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 5.160
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 5.161.
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 5.162
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 :

$$P_1 = X^4 + X + 1 \tag{5.328a}$$
$$P_2 = X^4 + X^3 + 1 \tag{5.328b}$$
$$P_3 = X^4 + X^3 + X^2 + X + 1. \tag{5.328c}$$

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

$$X \tag{5.329a}$$
$$X^2 \tag{5.329b}$$
$$X^3 \tag{5.329c}$$
$$X^4 = X^3 + X^2 + X + 1 \tag{5.329d}$$
$$1. \tag{5.329e}$$

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

$$X \tag{5.330a}$$
$$X^2 \tag{5.330b}$$
$$X^3 \tag{5.330c}$$
$$X+1 \tag{5.330d}$$
$$X^2+X \tag{5.330e}$$
$$X^3+X^2 \tag{5.330f}$$
$$X+1+X^3 \tag{5.330g}$$
$$X^2+1 \tag{5.330h}$$
$$X^3+X \tag{5.330i}$$
$$X+1+X^2 \tag{5.330j}$$
$$X^2+X+X^3 \tag{5.330k}$$
$$X^3+X^2+X+1 \tag{5.330l}$$
$$1+X^2+X^2 \tag{5.330m}$$
$$1+X^3 \tag{5.330n}$$
$$1 \tag{5.330o}$$

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 5.163.
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,

$$\mathbb{F}_q = \{0, \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{q-1}\}, \tag{5.331}$$

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 5.155. D'autre part le groupe $(\mathbb{F}_p[X]/P)^*$ est cyclique d'ordre $q-1$. Par conséquent le corollaire 3.55 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 5.132 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 $\mathbb{F}_q$.

□

5.7.7 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

$$P = X^4 + X^3 + 1 \tag{5.332}$$

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 :

$$\omega^5 = \omega^4\omega = (\omega^3 + 1)\omega = \omega^4 + \omega = \omega^3 + \omega + 1 \neq 1. \tag{5.333}$$

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

$$P(\beta^2) = (\beta^2)^4 + (\beta^2)^3 + 1 = (\beta^4)^2 + (\beta^3)^2 + 1^2 = (\beta^4 + \beta^3 + 1)^2 = 0. \tag{5.334}$$

Ici nous avons implicitement utilisé le lemme 5.150. D'autre part P ne peut pas avoir plus de 4 racines.

Proposition 5.164.
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

$$f_1 = \omega \tag{5.335a}$$
$$f_2 = \omega^2 \tag{5.335b}$$
$$f_3 = \omega^3 + 1 \tag{5.335c}$$
$$f_4 = \omega^3 + \omega^2 + \omega. \tag{5.335d}$$

Ensuite nous montrons que les vecteurs e_i peuvent être construits comme combinaisons linéaires des vecteurs f_j :

$$f_1 + f_2 + f_3 + f_4 = e_0 \tag{5.336a}$$
$$f_1 = e_1 \tag{5.336b}$$
$$f_2 = e_2 \tag{5.336c}$$
$$f_1 + f_2 + f_4 = e_3. \tag{5.336d}$$

Les quatre vecteurs f_j forment donc bien un base parce qu'ils sont générateurs d'un espace de dimension 4. □

Exemple 5.165

(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

$$\omega^{5k} = \omega^l \tag{5.337}$$

où l est donné. Cette équation revient à

$$5k = l \mod 15. \tag{5.338}$$

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 :

$$k = \begin{cases} 3, 6, 9, 12 & \text{si } l{=}0 \\ 1 & \text{si } l{=}5 \\ 2 & \text{si } l{=}10 \end{cases} \tag{5.339}$$

(2) Nous cherchons γ sous la forme $\gamma = \omega^k$. Parmi les nombreuses contraintes liées à l'énoncé nous devons avoir

$$\gamma^5 = 1, \gamma, \gamma^2, \gamma^4, \gamma^8. \tag{5.340}$$

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

$$\gamma = 1, \omega^3, \omega^6, \omega^9, \omega^{12}. \tag{5.341}$$

Évidemment $\gamma = 1$ ne produit pas une base. Avec $\gamma = \omega^3$ nous trouvons

$$B_\gamma = \{\omega^3, \omega^6, \omega^{12}, \omega^{24}\} = \{\omega^3, \omega^6, \omega^{12}, \omega^9\} \tag{5.342}$$

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

$$\omega^5 = \omega^3 + \omega + 1 \tag{5.343a}$$
$$\omega^6 = \omega^3 + \omega^2 + \omega + 1 \tag{5.343b}$$
$$\omega^9 = \omega^2 + 1 \tag{5.343c}$$
$$\omega^{12} = \omega + 1. \tag{5.343d}$$

L'ensemble B_γ est alors formé des éléments

$$f_1 = \omega^3 \tag{5.344a}$$
$$f_2 = \omega^3 + \omega^2 + \omega + 1 \tag{5.344b}$$
$$f_3 = \omega + 1 \tag{5.344c}$$
$$f_4 = \omega^2 + 1. \tag{5.344d}$$

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.

△

5.7.8 Polynômes irréductibles sur $\mathbb{F}_q$

Définition 5.166.
*La **fonction de Möbius** est la fonction $\mu\colon \mathbb{N}^* \to \{-1,0,1\}$ définie par*

$$\mu(n) = \begin{cases} 0 & \text{si } n \text{ est divisible par un carré différent de } 1, \\ 1 & \text{si } n \text{ est le produit d'un nombre pair de nombres premiers distincts,} \\ -1 & \text{si } n \text{ est le produit d'un nombre impair de nombres premiers distincts,} \end{cases} \tag{5.345}$$

Proposition 5.167 ([54])**.**
Si m et n sont strictement positifs et premiers entre eux, alors

$$\mu(mn) = \mu(m)\mu(n). \tag{5.346}$$

De plus nous avons

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n > 1 \end{cases} \tag{5.347}$$

Proposition 5.168 (Formule d'inversion de Möbius[54])**.**
Soient $f, g\colon \mathbb{N} \to \mathbb{C}$ telles que pour tout $n \geqslant 1$,

$$g(n) = \sum_{d|n} f(d). \tag{5.348}$$

Alors

$$f(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) g(d) \tag{5.349}$$

où μ est la fonction de Möbius pour tout $n \geqslant 1$.

Lemme 5.169 ([55])**.**
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[23] $aP + bQ = 1$. Cette dernière égalité est encore valable dans $\mathbb{L}$ et donc rend impossible l'existence d'une racine commune. □

Proposition 5.170 ([55, 2])**.**
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 :

$$X^{q^n} - X = \prod_{d|n} \prod_{p \in A(d,q)} P. \tag{5.350}$$

(2) Le nombre d'irréductibles est donné par

$$I(n,q) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) q^d \tag{5.351}$$

où μ est la fonction de Möbius (définition 5.166).

(3) Nous avons l'équivalence de suite

$$I(n,q) \sim_{n\to\infty} \frac{q^n}{n}. \tag{5.352}$$

23. Théorème de Bézout, 5.40.

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 5.158. 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 5.132.

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

$$\alpha^{q^n} = \alpha^{q^{kd}} = \alpha^{q^d q^{(k-1)d}} = \left(\alpha^{q^d}\right)^{q^{(k-1)d}} = \alpha^{q^{(k-1)d}}, \tag{5.353}$$

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 5.169 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 $X^{q^n}-X$:

$$\prod_{d\,|\,n}\prod_{P\in A(d,q)} P \mid X^{q^n}-X. \tag{5.354}$$

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 5.55) :

$$[\mathbb{F}_{q^n} : \mathbb{K}][\mathbb{K} : \mathbb{F}_q] = [\mathbb{F}_{q^n} : \mathbb{F}_q] = n, \tag{5.355}$$

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 5.132), 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 :

$$X^{q^n}-X \mid \prod_{d\,|\,n}\prod_{P\in A(d,q)} P. \tag{5.356}$$

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 :

$$q^n = \sum_{d\,|\,n} d\,\mathrm{Card}\left(A(d,q)\right) = \sum_{d\,|\,n} dI(d,q). \tag{5.357}$$

Nous pouvons utiliser la formule d'inversion de Möbius (proposition 5.168) pour les fonctions $g(n)=q^n$ et $f(n)=dI(n,q)$. Nous écrivons alors

$$f(n) = \sum_{d\,|\,n} \mu\left(\frac{n}{d}\right) q^d, \tag{5.358}$$

ou encore

$$I(n,q) = \frac{1}{n}\sum_{d\,|\,n} \mu\left(\frac{n}{d}\right) q^d, \tag{5.359}$$

ce qu'il fallait.

(3) Nous posons

$$r_n = \sum_{\substack{d|n \\ d<n}} \mu\left(\frac{n}{d}\right) q^d, \tag{5.360}$$

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[24] 1,

$$|r_n| \leqslant \sum_{d=1}^{\lfloor n/2 \rfloor} q^d = \frac{q - q^{\lfloor n/2 \rfloor}}{1-q} = q\frac{q^{\lfloor n/2 \rfloor} - 1}{q-1} \leqslant \frac{q^{\lfloor n/2 \rfloor + 1}}{q-1}. \tag{5.361}$$

D'autre part en reprenant la formule déjà prouvée,

$$I(n,q) = \frac{1}{n}\sum_{d|n} \mu\left(\frac{n}{d}\right) q^d = \frac{1}{n}\left(r_n + \mu\left(\frac{n}{n}\right) q^n\right) = \frac{r_n + q^n}{n}. \tag{5.362}$$

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 $n \to \infty$:

$$\frac{q^n + r_n}{n} \sim_{n\to\infty} \frac{q^n}{n}. \tag{5.363}$$

□

5.8 Constructions à la règle et au compas

Définition 5.171 ([56])**.**
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*

$$C(E) = \bigcup_{n=1}^{\infty} C_n(E). \tag{5.364}$$

*Un réel est **constructible** s'il est l'abscisse d'un point constructible.*

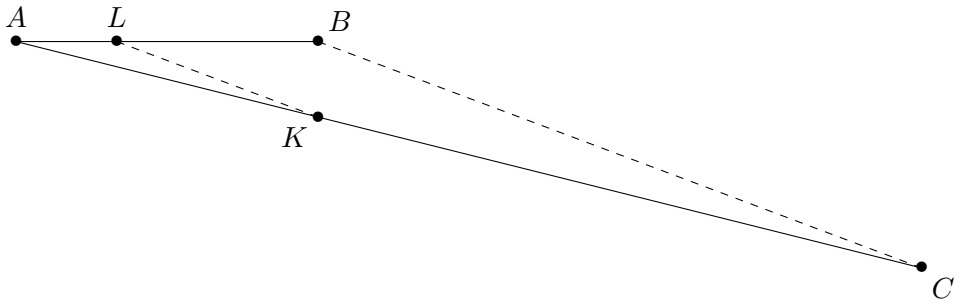
Pour toute la suite nous allons considérer les points et réels constructibles à partir de l'ensemble $E = \{(0,0),(0,1)\}$.

5.8.1 Quelque constructions

Proposition 5.172.
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 $n \in \mathbb{N}^*$.

24. Dans [2], ma dernière inégalité arrive comme une égalité.

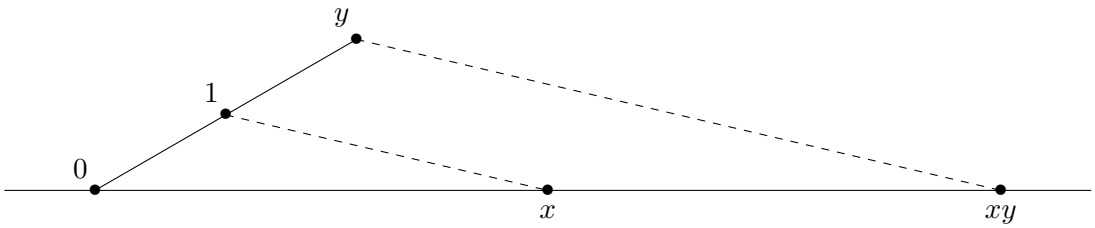


La méthode pour construire le nombre $1/n$ est la suivante. Soit $[AB]$ un segment de longueur 1 (par exemple $A = (0,0)$ et $B = (1,0)$) et un point C tel que $[AC]$ ait une longueur n. Nous plaçons sur $[AC]$ le point K situé à une distance 1 de A en pointant le compas en A et en traçant le cercle 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 5.173(Multiplication à la règle et au compas[56])
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. △

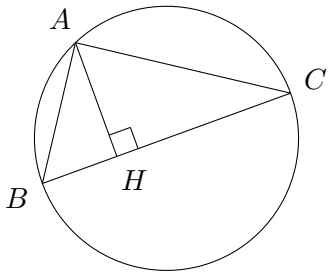
Exemple 5.174(Racine carré à la règle et au compas[56])

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 :

$$\frac{AB}{BH} = \frac{BC}{AB}, \tag{5.365}$$

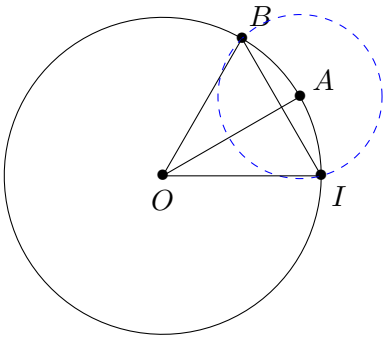
ce qui donne $AB^2 = BC \times BH = x$ (BC et BH valent respectivement 1 et x ou le contraire).



△

Exemple 5.175(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 5.176(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 $\beta + \alpha$. △

5.8.2 Nombres constructibles

Théorème 5.177 (Wantzel[57])**.**
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* $\mathbb{L}_i$

25. C'est à dire une extension finie de degré 2.

(3) $a \in \mathbb{L}_n$.

Démonstration. Soit $\mathbb{K}$ un corps de nombres constructibles (par exemple $\mathbb{Q}$) ; nous notons $E_{\mathbb{K}}$ l'ensemble des points de $\mathbb{R}^2$ dont les coordonnées sont dans $\mathbb{K}$. Ce sont des points forcément constructibles.

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

$$\begin{cases} ax + by + c = 0 & (5.366a) \\ a'x + b'y + c' = 0, & (5.366b) \end{cases}$$

dont les solutions sont encore dans $\mathbb{K}$.

Intersection droite-cerle L'équation d'un cercle est de la forme

$$(x - u)^2 + (y - v)^2 = r^2 \tag{5.367}$$

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

$$x^2 + y^2 + ux + vy + t = 0 \tag{5.368}$$

avec $u, v, t \in \mathbb{K}$. Il s'agit de voir où sont les solutions du système

$$\begin{cases} ax + by + c = 0 & (5.369a) \\ x^2 + y^2 + ux + vy + t = 0. & (5.369b) \end{cases}$$

Si $a \neq 0$ alors nous pouvons faire la substitution $x = -(c + by)/a$ et obtenir l'équation suivante pour y :

$$\left(\frac{-c - by}{a}\right)^2 + y^2 + u\left(\frac{-c - by}{a}\right) + vy + t = 0. \tag{5.370}$$

Cela est de la forme $P(y) = 0$ où P est un polynôme du second degré à coefficients dans $\mathbb{K}$.

Si $a = 0$ nous substituons y au lieu de x et le résultat est le même.

C'est le moment de relire la proposition 5.75 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 :

$$\big[\mathbb{K}[\alpha] : \mathbb{K}\big] = 2. \tag{5.371}$$

De plus $\mathbb{K}[\alpha] = \mathbb{K}(\alpha)$.

Intersection cercle-cercle Le système

$$\begin{cases} x^2 + y^2 + ux + vy + y = 0 & (5.372a) \\ x^2 + y^2 + ax + by + c = 0 & (5.372b) \end{cases}$$

est équivalent au système (en substituant la seconde équation par la différence entre les deux)

$$\begin{cases} x^2 + y^2 + ux + vy + y = 0 & (5.373a) \\ (u - a)x + (v - b)y + t - c = 0 & (5.373b) \end{cases}$$

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

$$a\alpha^2 + b\alpha + c = 0 \tag{5.374}$$

avec a, b, c trois réels constructibles en $m-1$ étapes. L'hypothèse de récurrence donne donc des piles d'extensions

$$\begin{array}{ccc} \mathbb{L}_0 = \mathbb{Q} & \mathbb{L}'_0 = \mathbb{Q} & \mathbb{L}''_0 = \mathbb{Q} \\ L_{i+1} = \mathbb{L}_i(a_i) & \mathbb{L}'_{i+1} = \mathbb{L}'_i(b_i) & \mathbb{L}''_{i+1} = \mathbb{L}''_i(c_i) \\ a \in \mathbb{L}_n & b \in \mathbb{L}'_{n'} & c \in \mathbb{L}''_{n''}. \end{array} \tag{5.375}$$

Nous considérons donc la suite d'extensions de $\mathbb{Q}$ qui consiste à étendre successivement par les nombres $a_1, \ldots, a_n, b_1, \ldots, b_{n'}, c_1, \ldots, c_{n''}$ tout en excluant les doublons : il est possible que par exemple b_3 soit déjà dans $\mathbb{Q}(a_1, \ldots, a_n)$. Cela nous fournit une suite d'extensions

$$\begin{array}{c} \mathbb{M}_0 = \mathbb{Q} \\ \mathbb{M}_{i+1} = \mathbb{M}_i(\alpha_i) \\ a, b, c \in \mathbb{M}_n. \end{array} \tag{5.376}$$

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 5.172.

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

$$a\alpha^2 + b\alpha + c = 0 \tag{5.377}$$

pour certains $a, b, c \in \mathbb{L}_i$ et donc

$$\alpha = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{5.378}$$

Les exemples 5.173 et 5.174 montrent que les produits et les racines carrées de nombres constructibles sont constructibles. Donc α est constructible.

□

5.8.3 Polygones constructibles

Définition 5.178.
*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 5.179 ([2])**.**
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 5.175.

Sens réciproque Le théorème de Bézout 3.32 nous donne $a, b \in \mathbb{Z}$ tels que $an + bm = 1$. Cela donne immédiatement

$$\frac{1}{mn} = \frac{a}{m} + \frac{b}{n}, \tag{5.379}$$

et donc

$$\frac{2\pi}{mn} = a\frac{2\pi}{m} + b\frac{2\pi}{n}, \tag{5.380}$$

ce qui fait que l'angle $2\pi/mn$ est une combinaison entière d'angles constructibles. Il est donc constructible par l'exemple 5.176.

□

Lemme 5.180.
Soit un entier $n \geqslant 3$ ayant

$$n = \prod_{i=1}^{k} p_i^{\alpha_i} \tag{5.381}$$

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

$$\frac{2\pi}{\prod_{i=1}^{k} p_i^{\alpha_i}} \tag{5.382}$$

est constructible. Par récurrence sur le lemme 5.179, cet angle est constructible si et seulement si les angles $\frac{2\pi}{p_i^{\alpha_i}}$ sont constructibles. □

Théorème 5.181 (Gauss-Wantzel[2])**.**
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 nombre 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 $\cos\big(2\pi/p^\alpha\big)$ l'est également et le théorème de Wantzel 5.177 nous indique que

$$\Big[\mathbb{Q}\Big(\cos(2\pi/p^\alpha)\Big) : \mathbb{Q}\Big] = 2^m \tag{5.383}$$

pour un certain m. Posons $q = p^\alpha$ et $\omega = e^{2i\pi/q}$. Grâce au corollaire 5.123, nous savons que le polynôme minimal de ω est le polynôme cyclotomique ϕ_q dont le degré est [26]

$$\varphi(q) = \varphi(p^\alpha) = p^{\alpha-1}(p-1). \tag{5.384}$$

Par conséquent

$$\big[\mathbb{Q}(\omega) : \mathbb{Q}\big] = p^{\alpha-1}(p-1). \tag{5.385}$$

Mais par ailleurs

$$\omega + \omega^{-1} = 2\cos(2\pi/q), \tag{5.386}$$

donc $\cos(2\pi/q) \in \mathbb{Q}(\omega)$. Et en multipliant (5.386) par ω nous trouvons le polynôme annulateur suivant pour ω :

$$\omega^2 - 2\cos\left(\frac{2\pi}{q}\right)\omega + 1 = 0. \tag{5.387}$$

Cela signifie que ω est de degré 2 dans $\mathbb{Q}\big(\cos(2\pi/q)\big)$, c'est à dire

$$\big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/q)\big)\big] = 2. \tag{5.388}$$

26. Voir juste en dessous de la définition 5.116.

En remettant bout à bout et en utilisant la propriété multiplicative des degrés des extensions [27],

$$\underbrace{\big[\mathbb{Q}(\omega):\mathbb{Q}\big]}_{p^{\alpha-1}(p-1)}=\underbrace{\Big[\mathbb{Q}(\omega):\mathbb{Q}\big(\cos(2\pi/p^\alpha)\big)\Big]}_{2}\underbrace{\Big[\mathbb{Q}\big(\cos(2\pi/p^\alpha)\big):\mathbb{Q}\Big]}_{2^m},\tag{5.389}$$

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 :

$$p=2^{\lambda 2^\beta}+1=\left(2^{2^\beta}\right)^\lambda+1,\tag{5.390}$$

mais le lemme 4.125(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

$$p=1+2^{2^\beta},\tag{5.391}$$

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

$$[\mathbb{Q}(\omega):\mathbb{Q}]=p-1.\tag{5.392}$$

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 5.75), 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

$$\phi_p\big(g(\omega)\big)=g\big(\phi_p(\omega)\big)=g(0)=0,\tag{5.393}$$

ce qui signifie que $g(\omega)$ est une racine du polynôme cyclotomique ϕ_p. Par définition 5.116, les racines sont $\{\omega,\omega^2,\dots,\omega^{p-1}\}$, ce qui signifie que l'action de g consiste à élever ω à une certaine puissance entre 1 et $p-1$. Le groupe G est donc d'ordre $|G|=p-1$ et a pour éléments

$$g_k\colon\omega\mapsto\phi^k\tag{5.394}$$

avec $k=1,\dots,p-1$. L'application [28]

$$\begin{aligned}\psi\colon G&\to(\mathbb{Z}/p\mathbb{Z})^*\\ g_k&\mapsto[k]_p.\end{aligned}\tag{5.395}$$

est un morphisme surjectif entre deux groupes finis de même cardinal, donc c'est un isomorphisme. Le corollaire 5.131 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 $0\leqslant i\leqslant n$:

$$\mathbb{K}_i=\{z\in\mathbb{Q}(\omega)\text{ tel que }g_0^{2^i}(z)=z\}.\tag{5.396}$$

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 $g_0^2(zz')=g_0^2(z)g_0^2(z')$.

27. Proposition 5.55.

28. Pour rappel, la notation $[k]_p$ est la classe de l'entier k modulo p, qui est un élément de $\mathbb{Z}/p\mathbb{Z}$.

$\underline{\mathbb{K}_n = \mathbb{Q}(\omega)}$ 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 $z \in \mathbb{Q}(\omega)$.

$\underline{\mathbb{K}_0 = \mathbb{Q}}$ 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 5.75(2) que $\{\omega^i\}_{i=0,\dots,p-2}$ est une base[29] de $\mathbb{Q}(\omega)$. Vu que $g_0 \colon \mathbb{Q}(\omega) \to \mathbb{Q}(\omega)$ est un isomorphisme, l'image d'une base est une base :

$$\{g_0(\omega)^i\}_{i=0,\dots,p-2} \tag{5.397}$$

est également une base de $\mathbb{Q}(\omega)$. Soit $z \in \mathbb{K}_0$ et décomposons-le dans cette base[30] :

$$z = \sum_{i=0}^{p-2} \lambda_i g_0(\omega)^i; \tag{5.398}$$

nous appliquons g_0 à cette égalité en tenant compte du fait que $g_0(z) = z$:

$$z = \sum_{i=0}^{p-2} \lambda_i g_0(\omega)^{i+1}. \tag{5.399}$$

En identifiant les coefficients (et en remarquant que le dernier terme dans (5.398) est le premier dans (5.399)) nous voyons que tous les coefficients sont égaux :

$$z = \lambda_0 g_0(\omega + \cdots + \omega^{p-1}) = \lambda_0 g_0\big(\phi_p(\omega) - 1\big) = -\lambda_0 \in \mathbb{Q}. \tag{5.400}$$

Dans cette chaîne d'égalités nous avons utilisé le fait que $\phi_p(X) = 1 + X + \cdots + X^{p-1}$ (corollaire 5.121) et que $\phi_p(\omega) = 0$ par définition.

$\underline{\mathbb{K}_i \subset \mathbb{K}_{i+1}}$ **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 :

$$g_0^{2^{i+1}}(z) = g_0^{2\times 2^i}(z) = (g_0^{2^i})^2(z) = z. \tag{5.401}$$

Afin de voir que l'inclusion est stricte nous montrons que l'élément

$$x = \sum_{k=0}^{2^{n-i}-1} g_0^{k2^{i+1}}(\omega) \tag{5.402}$$

est dans $\mathbb{K}_{i+1} \backslash \mathbb{K}_i$. Dans la base $\{1, \omega, \dots, \omega^{p-1}\}$ l'élément x a une composante ω avec le terme $k = 0$, mais si on lui applique $g_0^{2^i}$ nous obtenons

$$g_0^{2^i}(x) = \sum_{k=0}^{2^{n-i-1}-1} g_0^{2^i(1+2k)}(\omega). \tag{5.403}$$

Une composante ω pour cela demanderait d'avoir $2^i(1 + 2k) = \lambda 2^n$ avec $\lambda \in \mathbb{N}$; cela demanderait

$$k = \lambda 2^{n-i-1} - \frac{1}{2} \tag{5.404}$$

ce qui est impossible. Donc $x \notin \mathbb{K}_i$.

Nous montrons que $x \in \mathbb{K}_{i+1}$ de la même manière :

$$g_0^{2^{i+1}}(x) = \sum_{k=0}^{2^{n-i-1}-1} g^{2^{i+1}(k+1)}(\omega). \tag{5.405}$$

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 (5.402), et il se trouve dans (5.405) avec $k = 2^{n-i-1} - 1$. Donc $g_0^{2^{i+1}}(x) = x$ et $x \in \mathbb{K}_{i+1}$.

29. Ici [2] 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 $\mathbb{Q}$.

Les degrés dans la tour Par définition de n et de ω, et par la propriété multiplicative des degrés[31] nous avons :

$$2^n = p - 1 = \big[\mathbb{Q}(\omega) : \mathbb{Q}\big] = \underbrace{\big[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}\big] \dots \big[\mathbb{K}_1 : \mathbb{Q}\big]}_{n \text{ facteurs}}. \tag{5.406}$$

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

$$\big[\mathbb{K}_{n-1} : \mathbb{Q}\big] = 2^{n-1}. \tag{5.407}$$

$\cos(2\pi/p) \in \mathbb{K}_{n-1}$ 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 $f^2 = \text{Id}$ donc

$$\omega = f^2(\omega) = \omega^{2\lambda}, \tag{5.408}$$

ce qui donne $1 = \omega^{\lambda^2-1}$ ou encore que p divise $\lambda^2 - 1$ 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 $f \neq \text{Id}$ nous avons

$$[\lambda]_p = -[1]_p. \tag{5.409}$$

Cela fait $f(\omega) = \omega^{-1}$.

Par ailleurs $\mathbb{Q}(\omega)$ est un corps, donc il contient

$$\cos\left(\frac{2\pi}{p}\right) = \frac{1}{2}(\omega + \omega^{-1}). \tag{5.410}$$

Nous en déduisons que $\cos(2\pi/p)$ est un point fixe de $f = g_0^{2^{-1}}$:

$$f\big(\cos(2\pi/p)\big) = \frac{1}{2}\big(f(\omega) + f(\omega^{-1})\big) = \cos(2\pi/p). \tag{5.411}$$

Être un point fixe de $g_0^{2^{n-1}}$ signifie être un élément de $\mathbb{K}_{n-1}$:

$$\cos\left(\frac{2\pi}{p}\right) \in \mathbb{K}_{n-1}. \tag{5.412}$$

Questions de degrés Nous avons alors les (non)inclusions suivantes :

$$\mathbb{Q}\left(\cos(2\pi/p)\right) \subset \mathbb{K}_{n-1} \subsetneq \mathbb{K}_n = \mathbb{Q}(\omega), \tag{5.413}$$

ce qui fait que

$$1 < \big[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}\big] \leqslant \big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/p)\big)\big] = 2 \tag{5.414}$$

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 (5.388) (ici $\alpha = 1$). Les inégalités de (5.414) sont donc en réalité des égalités :

$$\big[\mathbb{Q}(\omega) : \mathbb{K}_{n-1}\big] = \big[\mathbb{Q}(\omega) : \mathbb{Q}\big(\cos(2\pi/p)\big)\big] = 2. \tag{5.415}$$

Cela, combiné au fait que $\mathbb{Q}\big(\cos(2\pi/p)\big) \subset \mathbb{K}_{n-1}$ donne

$$\mathbb{K}_{n-1} = \mathbb{Q}\big(\cos(2\pi/p)\big). \tag{5.416}$$

Mais nous savons déjà que $\big[\mathbb{K}_{n-1} : \mathbb{Q}\big] = 2^{n-1}$, donc

$$\big[\mathbb{Q}\big(\cos(2\pi/p)\big) : \mathbb{Q}\big] = 2^{n-1} \tag{5.417}$$

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 5.177.

31. Proposition 5.55

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 5.180), 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 5.182.
D'après Wikipédia[58], les seuls nombres de Fermat connus pour être premiers sont

$$F_0 = 3 \tag{5.418a}$$
$$F_1 = 5 \tag{5.418b}$$
$$F_2 = 17 \tag{5.418c}$$
$$F_3 = 257 \tag{5.418d}$$
$$F_4 = 65537. \tag{5.418e}$$

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 5.183.
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 $\frac{2\pi}{3}$. Voila qui règle un des vieux problèmes de l'antiquité.

5.8.4 Matrices

Proposition 5.184.
Nous avons

$$|\mathrm{GL}(n, \mathbb{F}_p)| = (p^n - 1)(p^n - p) \dots (p^n - p^{n-1}). \tag{5.419}$$

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. □

5.9 Minuscule morceau sur la théorie de Galois

Vous trouverez des détails et des preuves à propos de la théorie de Galois dans [59, 21].

Définition 5.185.
Soit $\mathbb{K}$, un corps.

*Le **groupe de Galois** d'une extension $\mathbb{L}$ de $\mathbb{K}$ est le groupe des automorphismes de $\mathbb{L}$ laissant $\mathbb{K}$ invariant.*

Le groupe de Galois d'un polynôme sur $\mathbb{K}$ est le groupe de Galois de son corps de décomposition sur $\mathbb{K}$.

Définition 5.186.
*Des éléments $b_1, \dots, b_n$ d'une extension de $\mathbb{K}$ sont **algébriquement indépendants** si ils ne satisfont à aucune relation du type*

$$\sum \alpha_{i_1 \dots i_n} b_1^{i_1} \dots b_n^{i_n} = 0 \tag{5.420}$$

avec $\alpha_{i_1 \dots i_n} \in \mathbb{K}$.

Nous disons que l'équation

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0 \tag{5.421}$$

est l'**équation générale** de degré n si les coefficients a_i sont algébriquement indépendants sur $\mathbb{K}$.

Théorème 5.187.
Le groupe de Galois d'un polynôme de degré n est isomorphe au groupe symétrique S_n.

Corollaire 5.188.
L'équation générale de degré n est résoluble par radicaux si et seulement si $n \leqslant 5$.

5.9.1 Théorème de d'Alembert

Théorème 5.189 (Théorème de d'Alembert[21, 60])**.**
Tout polynôme non constant, à coefficients complexes, admet au moins une racine complexe.

Corollaire 5.190.
Le corps $\mathbb{C}$ est algébriquement clos.

Corollaire 5.191 ([1])**.**
Tout polynôme de degré 3 à coefficients réels possède au moins une racine réelle.

Démonstration. Soient les racines λ_1, λ_2 et λ_3 du polynôme en question. Toutes trois sont dans $\mathbb{C}$. Supposons que λ_1 ne soit pas réelle. Alors λ_2 ou λ_3 doit être égale à $\bar{\lambda}_1$. Disons λ_2. Nous avons donc les racines λ_1, $\bar{\lambda}_1$ et λ_3. Le polynôme se factorise alors en

$$a(X - \lambda_1)(X - \bar{\lambda}_1)(X - \lambda_3). \tag{5.422}$$

Le coefficient a doit être réel parce qu'il est le coefficient du terme en X^3 (réel par hypothèse). Si λ_3 n'est pas réel, alors ce polynôme ne peut pas avoir des coefficients réels. Entre autres parce que terme indépendant est $a|\lambda_1|^2\lambda_3$, qui est réel si et seulement si λ_3 est réel [32]. □

Tant que vous y êtes, vous pouvez voir que le polynôme (5.422) est à coefficient réels si et seulement si $a \in \mathbb{R}$ et $\lambda_3 \in \mathbb{R}$.

32. Notez l'utilisation du lemme 2.55.

Chapitre 6

Topologie générale

6.1 Topologie en général

6.1.1 Définitions basiques

Définition 6.1.
Soit E, un ensemble et $\mathcal{T}$, une partie de l'ensemble de ses parties qui vérifie les propriétés suivantes

(1) les ensembles $\varnothing$ et E sont dans $\mathcal{T}$,

(2) Une union quelconque[1] *d'éléments de $\mathcal{T}$ est dans $\mathcal{T}$.*

(3) Une intersection finie *d'éléments de $\mathcal{T}$ est dans $\mathcal{T}$.*

Un tel choix $\mathcal{T}$ de sous-ensembles de E est une ***topologie*** *sur E, et les éléments de $\mathcal{T}$ sont appelés des* ***ouverts****. Nous disons que un sous ensemble A de E est* ***fermé*** *si son complémentaire, A^c est ouvert.*

Lemme 6.2.
Union et intersection de fermés.

(1) Une intersection quelconque de fermés est fermée.

(2) Une union finie de fermés est fermée.

Démonstration. Soit $\{F_i\}_{i\in I}$ est un ensemble de fermés ; nous avons

$$\left(\bigcap_{i\in I} F_i\right)^c = \bigcup_{i\in I} F_i^c. \tag{6.1}$$

L'union de droite est un ouvert (union d'ouvert), et donc le terme de gauche est le complémentaire d'un ouvert. Donc fermé.

Le complémentaire de l'union finie est une intersection finie de complémentaires et est donc ouvert. □

Dans un espace topologique, nous avons une caractérisation très importante des ouverts.

Théorème 6.3.
Une partie d'un espace topologique est ouverte si et seulement si elle contient un voisinage ouvert de chacun de ses éléments.

Démonstration. Soit X un espace topologique et $A \subset X$. Le sens direct est évident : A lui-même est un ouvert autour de $x \in A$, qui est inclus à A.

Pour le sens inverse, nous supposons que A contienne un ouvert atour de chacun de ses points. Pour chaque $x \in A$, nous considérons l'ensemble $\mathcal{O}_x \subset A$, un ouvert autour de x. Nous avons que

$$A = \bigcup_{x\in A} \mathcal{O}_x. \tag{6.2}$$

1. Par «quelconque» nous entendons vraiment quelconque : c'est à dire indicée par un ensemble qui peut autant être $\mathbb{N}$ que $\mathbb{R}$ qu'un ensemble encore considérablement plus grand.

En effet $A \subset \bigcup_{x\in A} \mathcal{O}_x$ parce que tous les éléments de A sont dans un des $\mathcal{O}_x$, par construction. D'autre part, $\bigcup_{x\in A} \mathcal{O}_x \subset A$ parce que chacun des $\mathcal{O}_x$ est compris dans A.

L'union du membre de droite de (6.2) est une union d'ouverts et est donc un ouvert. Cela prouve que A est un ouvert.

□

Une utilisation typique de ce théorème est faite dans le lemme 6.169.

Dès que nous avons une topologie nous avons une notion de convergence.

Définition 6.4 (Convergence de suite)**.**
*Une suite (x_n) d'éléments de E **converge** vers l'élément y de E si pour tout ouvert $\mathcal{O}$ contenant y, il existe un K tel que $k > K$ implique $x_k \in \mathcal{O}$.*

6.1.2 Topologie produit

Définition 6.5 (Produit d'espaces topologiques, thème 6)**.**
*Soient $X_1, \ldots, X_n$ des espaces topologiques. Leur **produit** est l'ensemble*

$$X = \prod_{i=1}^{n} X_i \tag{6.3}$$

muni de la topologie suivante. L'ensemble U est ouvert si et seulement si pour tout $x \in U$ il existe des ouverts $A_i \in X_i$ tels que $x \in A_1 \times \cdots \times A_n \subset U$.

Proposition 6.6 ([1])**.**
La convergence d'une suite pour la topologie de l'espace produit implique la convergence des suites «composantes par composante».

Démonstration. Pour simplifier les notations, nous allons considérer le produit de deux espaces. Soit donc $(x_k, y_k) \xrightarrow{X\times Y} (x, y)$ et des ouverts $\mathcal{O}_1$ dans X autour de x et $\mathcal{O}_2$ autour de y dans Y. La partie $\mathcal{O}_1 \times \mathcal{O}_2$ est ouverte dans $X \times Y$. Donc il existe K tel que $k > K$ implique $(x_k, y_k) \in \mathcal{O}_1 \times \mathcal{O}_2$.

Nous avons prouvé que pour tout ouvert $\mathcal{O}_1$ autour de x il existe K tel que $k > K$ implique $x_k \in \mathcal{O}_1$. Donc $x_k \xrightarrow{X} x$. Idem pour y. □

6.1.3 Séparabilité

Définition 6.7 (Espace séparé)**.**
*Si deux points distincts ont toujours deux voisinages distincts, nous disons que l'espace est **séparé** ou **Hausdorff**.*

La notion d'espace séparé est particulièrement importante parce qu'elle assure l'unicité de la limite d'une fonction en un point par la proposition 6.68.

Définition 6.8.
*Un espace topologique est **séparable** s'il possède une partie dénombrable dense.*

À ne pas confondre avec :

Définition 6.9.
*Un espace topologique est **séparé** ou **Hausdorff** si deux points distincts possèdent des voisinages disjoints.*

Définition 6.10.
*Une partie A d'un espace topologique est **compacte** s'il vérifie la propriété de Borel-Lebesgue : pour tout recouvrement de A par des ouverts (c'est-à-dire une collection d'ouverts dont la réunion contient A) on peut tirer un recouvrement fini.*

Remarque 6.11.
Certaines sources (dont wikipédia) disent que pour être compact il faut aussi être séparé [2]. Pour ces sources, un espace qui ne vérifie que la propriété de Borel-Lebesgue est alors dit **quasi-compact**.

Définition 6.12.
Une partie d'un espace topologique est ***relativement compact*** *si son adhérence est compacte.*

Définition 6.13.
Un espace topologique est ***localement compact*** *si tout élément possède un voisinage compact.*

Définition 6.14 (Séquentiellement compact)**.**
Nous disons qu'un espace topologique est ***séquentiellement compact*** *si toute suite admet une sous-suite convergente.*

Définition 6.15.
Un espace topologique est ***dénombrable à l'infini*** *s'il est réunion dénombrable de compacts.*

Définition 6.16 (Base de topologie)**.**
Une famille $\mathcal{B}$ d'ouverts de X est une ***base de la topologie*** *de X si pour tout $x \in X$ et pour tout voisinage V de x, il existe $A \in \mathcal{B}$ tel que $x \in A \subset V$.*

Proposition 6.17.
Si $\mathcal{B}$ est une base de la topologie de X alors tout ouvert de X est une union d'éléments de $\mathcal{B}$.

Démonstration. Soit $\mathcal{O}$ un ouvert de X ; pour chaque $x \in \mathcal{O}$ nous considérons un ouvert $U(x)$ tel que $x \in U(x) \subset \mathcal{O}$ (possible par le théorème 6.3). Nous prenons alors $B(x) \in \mathcal{B}$ tel que

$$x \in B(x) \subset U(x) \subset \mathcal{O}. \tag{6.4}$$

Alors nous avons $\mathcal{O} = \bigcup_{x\in\mathcal{O}} B(x)$. □

Notons toutefois que nous sommes loin d'avoir une union dénombrable en général.

6.1.4 Topologie métrique

Définition 6.18.
Si E est un ensemble, une ***distance*** *sur E est une application $d\colon E \times E \to \mathbb{R}$ telle que pour tout $x, y \in E$,*

(1) $d(x, y) \geqslant 0$
(2) $d(x, y) = 0$ *si et seulement si* $x = y$,
(3) $d(x, y) = d(y, x)$
(4) $d(x, y) \leqslant d(x, z) + d(z, y)$.

*La dernière condition est l'****inégalité triangulaire****.*
Un couple (E, d) formé d'un ensemble et d'une distance est un ***espace métrique****.*

La définition-théorème suivante donne une topologie sur les espaces métriques en partant des boules. Par construction, les boules sont une base de la topologie [3] d'un espace métrique.

Théorème-définition 6.19.
Soit (E, d) un espace métrique. Nous définissons les ***boules ouvertes*** *par*

$$B(x, r) = \{y \in E \text{ tel que } d(x, y) < r\}. \tag{6.5}$$

Alors en posant

$$\mathcal{T} = \big\{\mathcal{O} \subset E \text{ tel que } \forall x \in A, \exists r > 0 \text{ tel que } B(x, r) \subset \mathcal{O}\big\} \tag{6.6}$$

2. Définition 6.9.
3. Définition 6.16.

nous définissons une topologie sur E.

Cette topologie sur E est la ***topologie métrique*** *de (E,d). En présence d'une distance, sauf mention explicite du contraire, c'est toujours cette topologie-là que nous utiliserons.*

Démonstration. D'abord $\varnothing \in \mathcal{T}$ parce que tout élément de l'ensemble vide ... heu ... enfin parce que d'accord hein [4].

Ensuite si $(A_i)_{i\in I}$ sont des éléments de $\mathcal{T}$ et si $x \in \bigcup_{i\in I} A_i$ alors il existe $k \in I$ tel que $x \in A_k$. Par hypothèse il existe une boule $B(x,r) \subset A_k \subset \bigcup_{i\in I} A_i$.

Enfin si $(A_i)_{i\in\{1,\dots,n\}}$ sont des éléments de $\mathcal{T}$ alors pour tout i il existe $r_i > 0$ tel que $B(x,r_i) \subset A_i$. En prenant $r = \min\{r_i\}_{i=1,\dots,n}$ nous avons $B(x,r) \subset \bigcap_{i=1}^n A_i$. □

Remarque 6.20.
Deux remarques à propos de cette définition.

(1) Cette définition est faite exprès pour respecter le théorème 6.3.

(2) Par construction, les boules ouvertes sont une base de la topologie (définition 6.16) des espaces métriques.

Définition 6.21.
Un sous ensemble $A \subset \mathbb{R}^n$ est ***borné*** *s'il existe une boule de $\mathbb{R}^n$ contenant A.*

Proposition 6.22.
Toute réunion finie d'ensembles bornés est un ensemble borné. Toute partie d'un ensemble borné est un ensemble borné.

Définition 6.23.
Si (X,d_X) et (Y,d_Y) sont des espaces métrique, une ***isométrie*** *est une application bijective $f\colon X \to Y$ telle que pour tout $x,y \in X$ nous ayons*

$$d_Y\big(f(x),f(y)\big) = d_X(x,y). \tag{6.7}$$

Remarque 6.24.
Une application vérifiant (6.7) est automatiquement injective. En pratique, il ne faut donc vérifier que la surjectivité.

Exemple 6.25(Manque de surjectivité)
Si $X = [0,\infty[$ et $f(x) = x+1$ alors f vérifie (6.7) pour la distance $d(x,y) = |x-y|$, mais n'est pas surjective. △

Exemple 6.26
Si X est un ensemble, nous pouvons écrire la **métrique discrète**

$$d(x,y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{si } x \neq y. \end{cases} \tag{6.8}$$

Pour cette métrique, le groupe des isométries est le groupe symétrique de X, c'est à dire le groupe de toutes les bijections de X sur lui-même. △

Proposition 6.27.
Une isométrie est continue.

Proposition-définition 6.28 (Groupe des isométries)**.**
Si (X,d) est un espace métrique,

(1) l'ensemble des isométries de X, noté $\mathrm{Isom}(X)$ *est un groupe pour la composition.*

(2) Ce groupe agit fidèlement sur X.

4. Pour qui ne serait pas d'accord, allez ajouter $\varnothing$ dans la définition des ouverts et puis c'est tout.

6.1.5 Espace vectoriel topologique

Définition 6.29.
Un espace vectoriel V sur le corps $\mathbb{K}$ muni d'une topologie est un ***espace vectoriel topologique*** *si*

(1) La somme de deux vecteurs est une application continue $V \times V \to V$

(2) La multiplication par un scalaire est une application continue $\mathbb{K} \times V \to V$.

Note : le corps[5] *lui-même doit avoir sa topologie ; typiquement c'est $\mathbb{R}$ ou $\mathbb{C}$ muni de la topologie usuelle.*

Définition 6.30.
Une distance d sur un espace vectoriel topologique V est dite ***compatible*** *avec la topologie si la topologie induite*[6] *de d est celle de V.*

Une distance d sur un espace vectoriel V est dite ***invariante*** *si pour tout $x, y, u \in V$ nous avons*

$$d(x + u, y + u) = d(x, y). \tag{6.9}$$

Notons que lorsque nous parlons d'une distance compatible avec un espace vectoriel topologique, nous parlons de compatibilité avec la topologie, pas avec la structure vectorielle.

6.2 Espace vectoriel normé

6.2.1 Introduction : norme et valeur absolue

La valeur absolue est essentielle pour introduire les notions de limite et de continuité pour les fonctions d'une variable. En fait nous disons que la fonction $f : \mathbb{R} \to \mathbb{R}$ est continue au point a lorsque pour tout ε, il existe un δ tel que

$$|x - a| \leqslant \delta \Rightarrow |f(x) - f(a)| \leqslant \varepsilon. \tag{6.10}$$

La quantité $|x-a|$ donne la «distance» entre x et a ; la définition de la continuité signifie que pour tout ε, il existe un δ tel que si a et x sont au plus à la distance δ l'un de l'autre, alors $f(x)$ et $f(a)$ ne seront éloigné au plus d'une distance ε.

La valeur absolue, dans $\mathbb{R}$, nous sert donc à mesurer des distances entre les nombres. Les principales propriétés de la valeur absolue sont :

(1) $|x| = 0$ implique $x = 0$,

(2) $|\lambda x| = |\lambda||x|$,

(3) $|x + y| \leqslant |x| + |y|$

pour tout $x, y \in \mathbb{R}$ et $\lambda \in \mathbb{R}$.

Afin de donner une notion de limite pour les fonctions de plusieurs variables, nous devons trouver un moyen de définir les notion de «taille» d'un vecteur et de distance entre deux points de $\mathbb{R}^n$, avec $n > 1$. La notion de «taille» doit satisfaire propriétés analogues à celles de la valeur absolue.

La premier notion de «taille» pour un vecteur de $\mathbb{R}^2$ que nous vient à l'esprit est la longueur du segment entre l'origine et l'extrémité libre du vecteur. Cela peut être calculée à l'aide du théorème de Pythagore :

$$\text{taille de } (a, b) = \sqrt{a^2 + b^2}. \tag{6.11}$$

Nous pouvons introduire une la notion de distance entre les éléments de $\mathbb{R}^2$ de façon similaire :

$$d\big((a_x, a_y), (b_x, b_y)\big) = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2}. \tag{6.12}$$

Cette définition a l'air raisonnable ; est-elle mathématiquement correcte ? Peut-elle jouer le rôle de la valeur absolue dans $\mathbb{R}^2$? Est-elle la seule définitions possibles de «taille» et distance en $\mathbb{R}^2$?

5. Définition 2.21
6. Définition 6.19.

6.2.2 Norme

Nous voulons formaliser les notions de «taille» et de distance dans $\mathbb{R}^n$, et plus généralement dans un espace vectoriel V de dimension finie. Pour cela nous nous inspirons des propriétés de la valeur absolue.

Définition 6.31 ([61], thème 5)**.**
Soit E un espace vectoriel (pas spécialement de dimension finie) sur le corps $\mathbb{K}$ ($= \mathbb{R}$ ou $\mathbb{C}$). Une ***norme*** *sur E est une application $N : E \to \mathbb{R}^+$ telle que*

(1) $N(x) = 0$ si et seulement si $x = 0$;

(2) $N(\lambda x) = |\lambda| N(x)$ pour tout $\lambda \in \mathbb{R}$ et $x \in E$;

(3) $N(x + y) \leqslant N(x) + N(y)$

pour tout $x, y \in E$ et pour tout $\lambda \in \mathbb{K}$.

La propriété (3) est appelée ***inégalité triangulaire****.*

En prenant $\lambda = -1$ dans la propriété (2), nous trouvons immédiatement que $N(-x) = N(x)$.

Proposition 6.32.
Toute norme N sur l'espace vectoriel E vérifie l'inégalité

$$\big|N(x) - N(y)\big| \leqslant N(x - y) \tag{6.13}$$

pour tout $x, y \in E$.

Démonstration. Nous avons, en utilisant le point (3) de la définition 6.31,

$$N(x) = N(x - y + y) \leqslant N(x - y) + N(y), \tag{6.14a}$$
$$N(y) = N(y - x + x) \leqslant N(y - x) + N(x). \tag{6.14b}$$

Supposons d'abord que $N(x) \geqslant N(y)$. Dans ce cas, en utilisant (6.14a),

$$\big|N(x) - N(y)\big| = N(x) - N(y) \leqslant N(x - y) + N(y) - N(y) = N(x - y). \tag{6.15}$$

Si par contre $N(x) \leqslant N(y)$, alors nous utilisons (6.14b) et nous trouvons

$$\big|N(x) - N(y)\big| = N(y) - N(x) \leqslant N(y - x) + N(x) - N(x) = N(y - x). \tag{6.16}$$

Dans les deux cas, nous avons retrouvé l'inégalité annoncée. □

Cette proposition signifie aussi que

$$-N(x - y) \leqslant N(x) - N(y) \leqslant N(x - y). \tag{6.17}$$

Afin de suivre une notation proche de celle de la valeur absolue, à partir de maintenant, la norme d'un vecteur v sera notée $\|v\|$ au lieu de $N(v)$. La proposition 6.32 s'énoncera donc

$$\big|\|x\| - \|y\|\big| \leqslant \|x - y\|. \tag{6.18}$$

Définition 6.33.
Un espace vectoriel E muni d'une norme est une ***espace vectoriel normé****, et on écrit $(E, \|.\|)$.*

Lemme-définition 6.34 (Distance induite par une norme)**.**
Soit un espace vectoriel normé $(E, \|.\|)$. Nous posons

$$d(x, y) = \|x - y\|. \tag{6.19}$$

Alors

(1) d est invariante par translations : $d(a, b) = d(a + u, b + u)$

(2) d est une distance[7] *sur E.*

C'est la ***distance induite*** *par la norme.*

Démonstration. Le fait que la formule (6.19) soit invariante par translations est immédiat. En ce qui concerne le fait que ce soit une distance, le seul point délicat à vérifier est l'inégalité triangulaire. Nous avons $\|x\| = d(x, 0)$, donc nous pouvons écrire

$$d(x, y) = \|x - y\| \leqslant \|x\| + \|y\| = d(x, 0) + d(y, 0) = d(x, z) + d(y, z) \tag{6.20}$$

parce que la distance associée à la norme est invariante par translation. □

Un espace vectoriel normé est alors immédiatement un espace vectoriel topologique

6.2.3 Quelque exemples

Il est possible de définir de nombreuses normes sur $\mathbb{R}^n$. Citons en quelque unes.

Proposition-définition 6.35.
Les formules suivantes définissent des normes sur $\mathbb{R}^n$.

(1) Les normes $\|.\|_{L^p}$ ($p \in \mathbb{N}$) sont définies de la façon suivante :

$$\|x\|_{L^p} = \Big(\sum_{i=1}^{n} |x_i|^p \Big)^{1/p}, \tag{6.21}$$

pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

(2) La norme L^2 est la ***norme euclidienne.***

(3) Nous définissons également la ***norme supremum*** *par*

$$\|x\|_\infty = \max_i |x_i|. \tag{6.22}$$

Démonstration. Point par point[8] .

(1)

(2) Le fait que $x \mapsto \|x\|_{L^2}$ soit une norme provient de la propriété suivante :

$$\sqrt{(a + b)^2} \leqslant \sqrt{a^2} + \sqrt{b^2}, \tag{6.23}$$

laquelle se démontre en passant au carré :

$$(a + b^2) = a^2 + b^2 + 2ab \leqslant a^2 + b^2 + 2|ab| = \big(\sqrt{a^2} + \sqrt{b^2}\big)^2. \tag{6.24}$$

(3)

□

Parmi ces normes, celles qui seront le plus souvent utilisées dans ces notes sont

$$\begin{aligned} \|x\|_{L^1} &= \sum_{i=1}^{n} |x_i|, \\ \|x\|_{L^2} &= \Big(\sum_{i=1}^{n} |x_i|^2 \Big)^{1/2}. \end{aligned} \tag{6.25}$$

Soient $A = (A_x, A_y)$ et $B = (B_x, B_y)$ deux éléments de $\mathbb{R}^2$. La distance[9] euclidienne entre A et B est donnée par $\|A - B\|_2$. En effet, sur la figure 6.1, la distance entre les points A et B est donnée par

$$|AB|^2 = |AC|^2 + |CB|^2 = |A_x - B_x|^2 + |A_y - B_y|^2, \tag{6.26}$$

par conséquent,

$$|AB| = \sqrt{|A_x - B_x|^2 + |A_y - B_y|^2} = \|A - B\|_2. \tag{6.27}$$

7. Définition 6.18.
8. Preuve non terminée
9. Ne pas confondre «distance» et «norme».

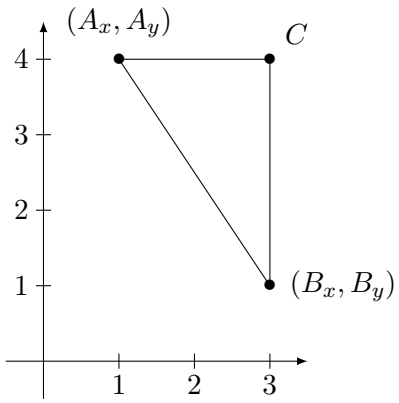


FIGURE 6.1 – La *norme* euclidienne induit la *distance* euclidienne. D'où son nom. Le point C est construit aux coordonnées (A_x, B_y).

Remarque 6.36.
Si A, B et C sont trois points dans le plan $\mathbb{R}^2$, alors l'inégalité triangulaire $|AB| \leqslant |AC| + |CB|$ est précisément la propriété (3) de la norme (définition 6.31). En effet l'inégalité triangulaire s'exprime de la façon suivante en terme de la norme $\|.\|_2$:

$$\|A - B\|_2 \leqslant \|A - C\|_2 + \|C - B\|_2. \tag{6.28}$$

En notant $u = A - C$ et $v = C - B$, l'équation (6.28) devient exactement la propriété de définition de la norme :

$$\|u + v\|_2 \leqslant \|u\|_2 + \|v\|_2. \tag{6.29}$$

Ceci explique pourquoi cette propriété des norme est appelée «inégalité triangulaire».

Les distances que nous avons vues jusqu'à présent sont des distances définies à partir d'une norme. La définition générale d'une distance est la définition 6.18.

6.3 Suites de Cauchy, métrique et espaces complet

6.3.1 Généralités

Définition 6.37 (Suite de τ-Cauchy, espace vectoriel topologique[62, 63])**.**
Soit E un espace vectoriel topologique. Une suite (x_k) dans E est une ***suite τ-Cauchy*** *si pour tout voisinage $\mathcal{U}$ de 0 il existe $N \in \mathbb{N}$ tel que $x_k - x_l \in \mathcal{U}$ pour tout $k, l \geqslant N$.*

Définition 6.38 (Espace τ-complet)**.**
Nous disons qu'une partie A d'un espace vectoriel topologique est ***τ-complet*** *si toute suite τ-Cauchy d'éléments de A converge vers un élément de A.*

Définition 6.39 (Suite de Cauchy, espace métrique)**.**
Une suite (a_k) dans un espace métrique (V, d) est ***de Cauchy*** *si pour tout ϵ, il existe N tel que si $n, m \geqslant N$ alors $d(a_n, a_m) < \epsilon$.*

Définition 6.40 (Métrique complete)**.**
Soit (E, d) un espace métrique. Nous disons que la métrique d est ***complète*** *si toute suite de Cauchy dans (E, d) converge dans E.*

6.41.
Ces définitions méritent quelque remarques.

(1) Dans le cas des espaces vectoriels topologiques, nous définissons les notions de suite τ-Cauchy et d'espace topologique τ-complet. Nous ajoutons le préfixe τ pour indiquer que ce sont des notions topologiques.

(2) Dans le cas des espaces métriques, nous définissons la notion de *métrique* complète. C'est bien la métrique qui est complète, et non l'espace. En effet nous allons voir dans l'exemple 6.42 que le même espace topologique peut accepter plusieurs distances différentes (donnant la même topologie) donnant lieu à des suites des Cauchy différentes.

(3) Si un espace vectoriel a une topologie issue d'une distance, rien ne dit que ses suites τ-Cauchy et ses suites de Cauchy sont les mêmes. Ce sont deux notions a priori séparées. Si V est un espace vectoriel topologique que l'on peut munir de deux distances d_1, d_2 donnant toutes deux la topologie, dire que V est τ-complet, dire que d_1 est complète et dire que d_2 est complète sont trois choses différentes. Même si les trois topologies sont identiques.

(4) Nous allons bien entendu voir que dans de larges gammes d'exemples, les notions de suite de Cauchy et τ-Cauchy coincident.

Exemple 6.42(La complétude n'est pas une propriété topologique[64])
Le fait pour un espace d'être complet n'est pas une propriété topologique, mais une propriété métrique. Plus exactement, il existe des espaces topologiques isomorphes, mais dont l'un est complet et l'autre non.

Nous considérons la distance suivante sur $\mathbb{N}$:

$$d_1(x,y) = |\frac{1}{x} - \frac{1}{y}|. \tag{6.30}$$

Pour vérifier que cette formule définit bien une distance (définition 6.18), le seul point non immédiat est l'inégalité triangulaire :

$$d_1(x,y) = |\frac{1}{x} - \frac{1}{y}| \leqslant |\frac{1}{x} - \frac{1}{z}| + |\frac{1}{z} - \frac{1}{y}| = d_1(x,z) + d_1(z,y). \tag{6.31}$$

Au niveau de la topologie induite par cette distance, c'est la topologie discrete. En effet, soit $x \in \mathbb{N}$ et $\epsilon > 0$; nous voulons déterminer la boule $B(x,\epsilon)$ en résolvant l'équation

$$|\frac{1}{x} - \frac{1}{y}| < \epsilon \tag{6.32}$$

pour $y \in \mathbb{N}$. Nous trouvons

$$\begin{cases} y > \dfrac{x}{\epsilon x + 1} & \text{(6.33a)} \\ y < \dfrac{x}{1 - \epsilon x}. & \text{(6.33b)} \end{cases}$$

Si x est assez petit, la seule solution entière est $y = x$. Les ouverts sont donc toutes les parties parce que tous les singletons sont ouvert.

Si d est la distance usuelle sur $\mathbb{N}$ ($d(x,y) = |x - y|$), nous avons donc isomorphisme d'espace topologique

$$(\mathbb{N}, d) \simeq (\mathbb{N}, d_1). \tag{6.34}$$

Nous pouvons même donner un isomorphisme explicite : $f(n) = n$.

La suite $(x_n) = n$ est une suite de Cauchy dans $(\mathbb{N}, d_1)$ parce que si $\epsilon > 0$ est donné, il suffit de prendre N assez grand pour avoir $\frac{1}{N} < \epsilon$ (possible par le lemme 2.52) nous avons, pour $n, m > N$:

$$|\frac{1}{n} - \frac{1}{m}| < \frac{1}{n} < \frac{1}{N} < \epsilon. \tag{6.35}$$

Or cette suite ne converge pas. Soit en effet un candidat limite k et calculons

$$d_1(x_n, k) = |\frac{1}{n} - \frac{1}{k}| \to \frac{1}{k} \neq 0. \tag{6.36}$$

L'espace $(\mathbb{N}, d_1)$ n'est pas complet.

Notons que cette suite n'est pas de Cauchy dans $(\mathbb{N}, d)$.

En résumé :

(1) Les espaces topologiques $(\mathbb{N}, d)$ et $(\mathbb{N}, d_1)$ sont isomorphes.
(2) Ils ont les mêmes notions de suites convergentes : une suite convergente pour l'un est convergente pour l'autre.
(3) Ils n'ont pas les mêmes notions de suites de Cauchy.
(4) Dans $(\mathbb{N}, d_1)$, il existe des suites de Cauchy qui ne convergent pas (pas complet).
(5) L'espace $(\mathbb{N}, d)$ est complet, mais $(\mathbb{N}, d_1)$ n'est pas complet.
(6) Le fait pour un espace topologique métrique d'être complet n'est pas intrinsèque à sa topologie : la complétude est une propriété de la distance. La complétude est une propriété de la métrique, et non de la topologie qui s'en suit.

△

6.3.2 Cas d'équivalence

Lemme 6.43.
Soit un espace vectoriel topologique V et une distance $d\colon V \times V \to \mathbb{R}^+$ compatible avec la topologie de V. Si d est invariante[10]*, alors les suites d-Cauchy et les suites τ-Cauchy sont les mêmes.*

Démonstration. Nous avons deux implications à prouver.

d-Cauchy implique τ-Cauchy Soit (x_n), une suite d-Cauchy dans V, et un voisinage U de 0. Vu que d est compatible avec la topologie de V, il existe une boule ouverte $B(0,\epsilon)$ inclue à U. Soit $N > 0$ tel que $m, n > N$ implique $d(x_n, x_m) < \epsilon$. Par invariance de la métrique, nous avons aussi

$$d(0, x_m - x_n) < \epsilon, \tag{6.37}$$

c'est à dire $x_m - x_n \in B(0,\epsilon) \subset U$. La suite (x_n) est donc τ-Cauchy.

τ-Cauchy implique d-Cauchy Soit (x_n), une suite τ-Cauchy dans V et $\epsilon > 0$. Vu que $B(0,\epsilon)$ est un voisinage de 0 dans V, il existe N tel que $m, n > N$ implique $x_n - x_m \in B(0,\epsilon)$. Cela signifie que $d(0, x_n - x_m) < \epsilon$ et toujours par invariance, que $d(x_n, x_m) < \epsilon$.

□

Définition 6.44 (Espace vectoriel topologique métrisable[65])**.**
*Un espace vectoriel topologique est **métrisable** si il existe une distance compatible avec la topologie.*

Proposition 6.45 ([66])**.**
Soit un espace topologique métrisable X.

(1) Tout fermé de X est une intersection dénombrable d'ouverts.
(2) Tout ouvert de X est une union dénombrable de fermés.

Démonstration. Soit une métrique d compatible avec la topologie de X et un fermé A. Nous posons

$$V_n = \{x \in X \text{ tel que } d(x, A) < \frac{1}{n}\}. \tag{6.38}$$

Et juste pour faire simple nous notons $V_0 = X$.

Les parties V_n sont ouvertes Soit $x \in V_n$ et trouvons un voisinage de x contenu dans V_n – encore le théorème 6.3. Si $y \in B(x,\epsilon)$ alors il existe $a \in A$ tel que $d(x,a) < \frac{1}{n}$ (ici les inégalités strictes sont importantes) et donc

$$d(y,a) \leqslant d(y,x) + d(x,a) < \epsilon + \frac{1}{n}. \tag{6.39}$$

Nous pouvons donc choisir ϵ de telle sorte que $d(y,a) < \epsilon$ pour tout $a \in B(x,\epsilon)$. Cela prouve que V_n est ouvert.

10. Définition 6.30.

A est l'intersection des V_n Nous avons évidemment $A \subset V_n$ pour tout n. Et d'autre part, si $a \in \bigcap_{n\in\mathbb{N}} V_n$ alors $d(a, A) < \frac{1}{n}$ pour tout n. Cela implique $d(a, A) = 0$, c'est à dire $a \in A$.

En ce qui concerne la seconde partie, nous passons au complémentaire. Si $\mathcal{O}$ est ouvert, $\mathcal{O}^c$ est fermé et

$$\mathcal{O}^c = \bigcap_{n\in\mathbb{N}} V_n, \tag{6.40}$$

ce qui donne immédiatement

$$\mathcal{O} = \bigcup_{n\in\mathbb{N}} V_n^c \tag{6.41}$$

où les V_n^c sont fermés. □

Corollaire 6.46.
Si X est un espace topologique métrisable, alors X accepte une base dénombrable de topologie autour de chaque point.

Démonstration. Il s'agit seulement de remarquer que les singletons sont fermés et d'appliquer la proposition 6.45. □

Théorème 6.47 ([63])**.**
Si V est un espace vectoriel topologique possédant en tout point une base de topologie dénombrable, alors il existe une distance d sur V telle que

(1) d est compatible avec la topologie de V,

(2) d est invariante par translation.

Une preuve est donnée dans [63] et je vous préviens : c'est pas simple.

Proposition 6.48.
Un espace vectoriel topologique est métrisable si et seulement si il possède en tout point une base dénombrable de topologie.

Démonstration. Il s'agit seulement de mettre bout à bout les corollaires 6.46 et théorème 6.47. □

Tout ceci nous mène à donner une large classe d'espaces vectoriels topologiques sur lesquelles les notions de suites d-Cauchy et τ-Cauchy coïncident.

Théorème 6.49.
Soit V un espace vectoriel topologique métrisable[11]*, alors il admet une métrique d compatible avec la topologie telle que une suite dans V est d-Cauchy si et seulement si elle est τ-Cauchy.*

Démonstration. Soit d une métrique sur V satisfaisant au théorème 6.47. Vu qu'elle est invariante par translation, les suites d-Cauchy sont exactement les suites τ-Cauchy par le lemme 6.43. □

Remarque 6.50.
Même si V est métrisable, si on choisit la métrique n'importe comment, on ne peut rien espérer.

6.51.
Sur les espaces vectoriels topologiques métrisables, nous pouvons donc parler de suite de Cauchy sans préciser si nous parlons de τ-Cauchy ou de d-Cauchy, parce que nous sous-entendons avoir choisi une métrique non seulement compatible avec la topologie, mais également invariante par translation.

Il reste cependant à traiter le cas d'un espace vectoriel topologique non métrisable. Dans ce cas, il n'y a pas de métrique, et la question de l'équivalence des définitions ne se pose pas.

11. i.e. admet une base dénombrable de topologique, voir la proposition 6.48

6.3.3 Espace topologique normé

Dans les espaces topologiques normés, il n'y a pas d'ambiguïté.

Proposition 6.52.
Soit (E, d) un espace vectoriel topologique normé.

(1) Une suite (x_n) dans E est convergente[12] *vers x si et seulement si pour tout $\epsilon \in \mathbb{R}$ il existe N_ϵ tel que pour tout $n \geqslant N_\epsilon$ nous avons $d(x_n, x) \leqslant \epsilon$.*

(2) Une suite (x_n) dans E est de Cauchy[13] *si pour tout $\epsilon \in \mathbb{R}$, il existe un N_ϵ tel que si $p, q \geqslant N_\epsilon$, nous avons $d(x_p, x_q) \leqslant \epsilon$.*

Démonstration. En ce qui concerne la convergence :

Sens direct Nous supposons que $x_k \to x$ dans E. Soit $\epsilon > 0$; vu que $B(x, \epsilon)$ est un ouvert contenant x, il existe un $N_\epsilon > 0$ tel que $k > N_\epsilon$ implique $x_k \in B(x, \epsilon)$. Cela signifie $d(x, x_k) \leqslant \epsilon$.

Réciproque Nous supposons que pour tout $\epsilon > 0$, il existe $N_\epsilon > 0$ tel que si $k > N_\epsilon$ alors $x_k \in B(x, \epsilon)$. Soit un ouvert $\mathcal{O}$ autour de x. Nous sommes dans un espace métrique ; ergo la topologie est donné par le théorème 6.19 et en particulier la liste des ouverts est donnée par (6.6). Il existe donc une boule $B(x, \epsilon)$ inclue à $\mathcal{O}$. Pour tout $k > N_\epsilon$ nous avons alors $x_k \in B(x, \epsilon) \subset \mathcal{O}$.

En ce qui concerne les suites de Cauchy :

Sens direct Si (x_n) est une suite de Cauchy et si $\epsilon > 0$ est donné, alors $B(0, \epsilon)$ est un voisinage de 0 et il existe N_ϵ tel que si $p, q \geqslant N_\epsilon$ alors $x_p - x_q \in B(0, \epsilon)$. Posons $u = x_p - x_q$; en utilisant l'invariance par translation (lemme 6.34(1).) nous avons

$$d(u, 0) = d(x_p - x_q, 0) = d(x_p, x_q). \tag{6.42}$$

Par conséquent $d(x_p, x_q) \leqslant \epsilon$.

Réciproque Soit $\mathcal{O}$ un voisinage de 0. Il existe ϵ tel que $B(0, \epsilon) \subset \mathcal{O}$. Par hypothèse il existe N_ϵ tel que $d(x_p, x_q) \leqslant \epsilon$ dès que $p, q \geqslant N_\epsilon$. En utilisant encore l'invariance par translation nous avons

$$d(x_p, x_q) = d(x_p - x_q, 0), \tag{6.43}$$

et comme cela est plus petit que ϵ, nous avons $x_p - x_q \in B(0, \epsilon) \subset \mathcal{O}$.

□

6.3.4 Espace métrique

Proposition 6.53 ([67])**.**
Toute suite convergente dans un espace métrique est de Cauchy.

Démonstration. Nous utilisons les caractérisations de la proposition 6.52 des suites convergentes et de Cauchy.

Soit un espace métrique (X, d) et $x_n \to \ell$ une suite convergente. Si $\epsilon > 0$, la proposition 6.52(1), dit qu'il existe N tel que pour tout $n > N$ nous ayons $d(x_n, \ell) < \epsilon$. Par conséquent si $n, m > N$ alors

$$d(x_n, x_m) \leqslant d(x_m, \ell) + d(l, x_m) \leqslant 2\epsilon. \tag{6.44}$$

Cela prouve que (x_n) est de Cauchy. □

12. Définition 6.4.
13. Définition 6.37.

6.4 Topologie sur l'ensemble des réels

Nous allons à présent donner la topologie sur $\mathbb{R}$ et ainsi résoudre les questions laissées en suspend lors de la construction des réels, voir 2.37.

Lemme 6.54.
L'application

$$x \mapsto |x| \tag{6.45}$$

est une norme sur $\mathbb{R}$.

La valeur absolue est celle de la définition 2.25(2).

Démonstration. Les conditions de la définition 6.31 sont immédiatement vérifiées grâce au lemme 2.27 et à la remarque 2.28. □

6.55.
Tant sur $\mathbb{Q}$ que sur $\mathbb{R}$, nous considérons la topologie métrique correspondant à cette norme (hors cas rarissimes qui seront signalés). Nous utiliserons toujours les caractérisations de la proposition 6.52 pour parler de suites convergentes et de suites de Cauchy.

Lemme 6.56 ([8])**.**
Toute suite de Cauchy dans $\mathbb{Q}$ *converge dans* $\mathbb{R}$ *vers le réel qu'elle représente.*

Démonstration. Soit (x_n) une suite de Cauchy de $\mathbb{Q}$, c'est à dire que $x_k \in \mathbb{Q}$ pour tout k et qu'elle est de Cauchy. Elle représente un réel $\bar{x} \in \mathbb{R}$, et nous voulons prouver que pour la topologie de $\mathbb{R}$ nous avons $\lim_{n\to\infty} x_n = \bar{x}$. Dans cette dernière limite, chacun des x_n est vu dans $\mathbb{R}$.

Si $\epsilon \in \mathbb{Q}$ est donné, il existe N_ϵ tel que si $p, q \geqslant N_\epsilon$ alors $|x_p - x_q| < \epsilon$, c'est à dire

$$x_p - \epsilon < x_q < x_p + \epsilon. \tag{6.46}$$

Soit $p \geqslant N_\epsilon$ fixé. Pour tout $q \geqslant N_\epsilon$ nous avons $x_p - \epsilon < x_q < x_p + \epsilon$. Par conséquent, la suite $n \mapsto (x_p - \epsilon) - q_n$ est un élément de $\mathcal{E}^-$ et au niveau des classes nous pouvons écrire

$$\overline{n \mapsto (x_p - \epsilon) - x_n} \leqslant 0. \tag{6.47}$$

Vu que $x_p - \epsilon$ représente la suite constante nous avons l'inégalité suivante dans $\mathbb{R}$:

$$x_p - \epsilon - \bar{x} \leqslant 0 \tag{6.48}$$

ou encore : $\bar{x} \geqslant x_p - \epsilon$. En faisant de même avec l'autre partie de l'inégalité, $x_p - \epsilon \leqslant \bar{x} \leqslant x_p + \epsilon$, ce qui signifie que

$$x_p \in B(\bar{x}, \epsilon) \tag{6.49}$$

dès que $p \geqslant N_\epsilon$. Cela signifie que $x_p \to \bar{x}$ dans $\mathbb{R}$. □

Théorème 6.57 (Complétude de $\mathbb{R}$[8])**.**
Nous avons :

(1) L'espace topologique $(\mathbb{R}, d)$ *est complet.*

(2) Une suite dans $\mathbb{R}$ *est convergente si et seulement si elle est de Cauchy.*

Démonstration. Soit (x_n) une suite de Cauchy dans $\mathbb{R}$. Pour chaque n, il existe par le lemme 6.56 un $y_n \in \mathbb{Q}$ tel que

$$x_n - \frac{1}{n} < y_n < x_n + \frac{1}{n}. \tag{6.50}$$

Nous allons prouver que (y_n) est une suite de Cauchy [14] dans $\mathbb{Q}$. Vu que (x_n) est de Cauchy pour la topologie de $\mathbb{R}$ si $\epsilon > 0$ dans $\mathbb{R}$ est donné, il existe n_ϵ tel que si $p, q \geqslant n_\epsilon$, alors [15] $|x_p - x_q| < \epsilon$.

Nous avons :

$$|y_p - y_q| \leqslant |y_p - x_p| + |x_p - x_q| + |x_q - y_q| < \frac{1}{p} + \epsilon + \frac{1}{q}. \tag{6.51}$$

En choisissant $N_\epsilon > \max\{n_\epsilon, \frac{1}{\epsilon}\}$ (ce qui est possible par le lemme 2.49), et en prenant $p, q > N_\epsilon$, nous avons

$$|y_p - y_q| \leqslant 3\epsilon, \tag{6.52}$$

ce qui prouve que (y_p) est une suite de Cauchy dans $\mathbb{Q}$ [16]. Elle représente donc un réel $\bar{y}$. Nous allons prouver que $x_n \xrightarrow{\mathbb{R}} \bar{y}$.

Nous savons qu'une suite de Cauchy de rationnels converge dans $\mathbb{R}$ vers le réel qu'elle représente, c'est à dire : $y_n \xrightarrow{\mathbb{R}} \bar{y}$ où chaque $y_n \in \mathbb{Q}$ est vu comme la suite constante (cela est le lemme 6.56). Autrement dit, pour $\epsilon > 0$, il existe un $N_\epsilon \in \mathbb{N}$ tel que si $p > N_\epsilon$ alors $|\bar{y} - y_p| < \epsilon$. Pour un tel p nous avons

$$|\bar{y} - x_p| \leqslant |\bar{y} - y_p| + |y_p - x_p| \leqslant \epsilon + \frac{1}{p}. \tag{6.53}$$

Donc dès que p est plus grand que $\max\{N_\epsilon, \frac{1}{\epsilon}\}$, nous avons $|\bar{y} - x_p| < 2\epsilon$, ce qui signifie que la suite (x_n) converge vers $\bar{y}$ dans $\mathbb{R}$.

Ceci achève de prouver que $\mathbb{R}$ est complet.

En ce qui concerne l'affirmation suivante, si une suite est convergente, alors elle est de Cauchy par la proposition 2.30(1). Et si une suite est de Cauchy, alors elle est convergente par ce que nous venons de prouver à propos de la complétude de $\mathbb{R}$. □

Proposition 6.58.
Les rationnels sont denses dans les réels.

Démonstration. Soit $r \in \mathbb{R}$ et $\epsilon \in \mathbb{R}^+$. Nous devons prouver l'existence d'un rationnel dans $B(x, \epsilon)$. Le lemme 2.52 dit qu'il existe un rationnel dans $]x - \epsilon/2, x + \epsilon/2[$ et donc dans $B(x, \epsilon)$. □

Proposition 6.59 ([1])**.**
Quel que soit le réel r, il existe une suite croissante de rationnels convergente vers r.

Démonstration. Soit $x \in \mathbb{R}$ et $\delta \in \mathbb{R}$; vu que $x - \delta$ et x sont des réels, le lemme 2.52 donne un élément x_δ tel que

$$x - 2\delta < x_\delta < x. \tag{6.54}$$

Il ne reste plus qu'à pêcher parmi ces x_δ pour trouver une suite croissante. Soit x_0 un rationnel plus petit que x. Nous posons $\delta_0 = x - x_0$ et ensuite :

$$\begin{cases} \delta_i = x - x_i & \text{(6.55a)} \\ x_{i+1} = x_{\delta_i/2}. & \text{(6.55b)} \end{cases}$$

Ainsi nous avons pour tout i les inégalités

$$x_i = x - \delta_i < x - \frac{\delta_i}{2} < x_{i+1} < x. \tag{6.56}$$

14. Soit au sens de la définition 2.25(3), soit au sens de la définition 6.37 (précisée par la proposition 6.52). Dans le premier cas, il faudra prendre $\epsilon \in \mathbb{Q}$ au lieu de $\epsilon \in \mathbb{R}$, et nous prouvons que $\mathbb{R}$ est un corps complet au lieu d'un espace topologique complet. Mais ça ne change pas grand chose.

15. Grâce à la proposition 6.52.

16. Éventuellement on peut préciser «pour la notion de suite de Cauchy de $\mathbb{Q}$», si on veut prouver que $\mathbb{R}$ est seulement un corps complet.

La suite est donc croissante et toujours plus petite que x. Mais nous avons à chaque étape $\delta_{i+1} < \frac{\delta_i}{2}$, ce qui donne $\delta_i \to 0$. Soit $\epsilon > 0$ et k tel que $\delta_k < \epsilon$. Nous avons alors

$$x_{k+1} \in B(x, \frac{\delta_k}{2}) \subset B(x, \epsilon). \tag{6.57}$$

□

6.5 Base de topologie

Proposition 6.60.
Un espace métrique séparable[17] *accepte une base de topologie dénombrable.*

Soit A dense et dénombrable dans l'espace métrique séparable (E, d). Si $\{a_i\}_{i\in\mathbb{N}}$ est une énumération de A et $\{r_i\}_{i\in\mathbb{N}}$ une énumération de $\mathbb{Q}$, alors

$$\mathcal{B} = \{B(a_i, r_j)\}_{i,j\in\mathbb{N}} \tag{6.58}$$

est une base de la topologie de E.

Démonstration. Soit $x \in E$ et V un voisinage de x. Ce dernier contient une boule $B(x, r)$ et quitte à prendre r un peu plus petit nous supposons que $r \in \mathbb{Q}$ (densité de $\mathbb{Q}$ dans $\mathbb{R}$, proposition 6.58).

Soit $a \in A$ avec $\|a - x\| < \frac{r}{3}$ (existe par densité de A dans E) ; nous avons $B(a, \frac{2r}{3}) \subset B(x, r)$ parce que si $y \in B(a, \frac{2r}{3})$ alors

$$\|y - x\| \leqslant \|y - a\| + \|a - x\| < \frac{2}{3}r + \frac{1}{3}r = r. \tag{6.59}$$

La seconde inégalité est stricte parce que les boules sont ouvertes. Le tout montre que $y \in B(x, r)$. Par ailleurs $x \in B(a, \frac{2}{3}r)$ et nous avons trouvé un élément de $\mathcal{B}$ contenant x tout en étant inclus à V. Cela prouve que $\mathcal{B}$ est bien une base de la topologie de E. □

Remarque 6.61.
Il est vite vu que les cubes ouverts forment aussi une base de la topologie de $\mathbb{R}^n$. Cela est à mettre en rapport avec le fait que toutes les normes sont équivalentes sur $\mathbb{R}^n$ (proposition 10.3).

Voir aussi le corollaire 13.192 qui donnera tout ouvert comme union de pavés presque disjoints.

Lemme 6.62.
Soient (X_1, d_1) et (X_2, d_2) des espace métriques séparables. Alors $X_1 \times X_2$ admet une base dénombrable de topologie constituée de produits de boule de X_1 par des boules de X_2. Plus précisément si A_i est dénombrable et dense dans X_i alors l'ensemble des produits

$$\big\{B(y_1, r_1) \times B(y_2, r_2)\big\}_{\substack{y_i\in A_i \\ r_i\in\mathbb{Q}^+}} \tag{6.60}$$

est une base de topologie pour $X_1 \times X_2$.

Démonstration. Soit $\mathcal{O}$ un ouvert de $X_1 \times X_2$ et $(x_1, x_2) \in \mathcal{O}$. Par définition de la topologie produit[18], il existe $r_1, r_2 \in \mathbb{Q}^+$ tels que $B(x_1, r_1) \times B(x_2, r_2) \subset \mathcal{O}$. Les parties A_i étant denses, il existe $y_i \in B(x_i, r_i/2) \cap A_i$. Avec ces choix nous avons $x_i \in B(y_i, \frac{r_i}{2})$. Nous avons donc

$$(x_1, x_2) \in B(y_1, \frac{r_1}{2}) \times B(y_2, \frac{r_2}{2}). \tag{6.61}$$

Il est facile de voir que $B(y_i, r_i/2) \subset B(x_i, r_i)$. En effet si $z_i \in B(y_i, r_i/2)$ alors

$$d_i(z_i, x_i) \leqslant d(z_i, y_i) + d(y_i, x_i) \leqslant \frac{r_i}{2} + \frac{r_i}{2} = r_i. \tag{6.62}$$

17. Qui possède une partie dense dénombrable, définition 6.8.
18. Définition 6.5.

Au final,

$$(x_1, x_2) \in B(y_1, \frac{r_1}{2}) \times B(y_2, \frac{r_2}{2}) \subset \mathcal{O}. \tag{6.63}$$

□

Lemme 6.63.
Une partie K d'un espace topologique est compacte si et seulement si de tout recouvrement par des ouverts d'une base de topologie nous pouvons extraire un sous-recouvrement fini.

Démonstration. Soit K une partie d'un espace topologique et $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de K par des ouverts. Chacun des $\mathcal{O}_i$ est une union d'éléments de la base de topologie par la proposition 6.17. Soient $\{A_j\}_{j\in J}$ un ensemble d'éléments de la base de topologie tel que chacun des $\mathcal{O}_i$ est une union de certains des A_j. Nous avons $\bigcup_{j\in J} A_j = \bigcup_{i\in I} \mathcal{O}_i$.

Par hypothèse nous pouvons extraire un ensemble fini $J_0 \subset J$ tel que $K \subset \bigcup_{j\in J_0} A_j$. Par construction chacun des A_j est inclus à (au moins) un des $\mathcal{O}_i$. Le choix d'un élément de I pour chacun des éléments de J_0 donne une partie finie I_0 de I telle que $K \subset \bigcup_{j\in J_0} A_j \subset \bigcup_{i\in I_0} \mathcal{O}_i$. □

6.6 Limite et continuité

Définition 6.64.
*Un **homéomorphisme** est une application bijective continue entre deux espaces topologiques dont la réciproque est continue. Deux espaces topologiques homéomorphes sont dits **isomorphes**.*

Définition 6.65 (Limite d'une fonction)**.**
*Soient X et Y des espaces topologiques, $A \subset X$ et $a \in \bar{A}$. L'élément $y \in Y$ est une **limite** de $f\colon A \to Y$ en a si pour tout voisinage V de y, il existe un voisinage W de a dans X tel que*

$$f(W \cap A) \subset V. \tag{6.64}$$

*s'il y a unicité de l'élément vers lequel f peut converger, alors nous disons que cet élément est la **limite** de f et nous notons*

$$\lim_{x\to a} f(x) = y. \tag{6.65}$$

Remarque 6.66.
La proposition 6.198 nous dira que l'unicité est de mise dans le cas des espaces duaux pour la topologie $*$-faible. La proposition 6.68 nous dira qu'il y a unicité dès que l'espace d'arrivée est séparé.

Exemple 6.67
Oui, il y a moyen de converger vers plusieurs points distincts si l'espace n'est pas super cool. Nous pouvons par exemple[68] considérer la droite réelle munie de sa topologie usuelle et y ajouter un point $0'$ (qui clone le réel 0) dont les voisinages sont les voisinages de 0 dans lesquels nous remplaçons 0 par $0'$. Dans cet espace, la suite $(1/n)$ converge à la fois vers 0 et $0'$. △

Proposition 6.68 (Unicité de la limite pour un espace séparé)**.**
Soient X un espace topologique, A une partie de X et Y un espace topologique séparé[19]*. Nous considérons une fonction $f\colon A \to Y$. Si $a \in \bar{A}$, alors f admet au plus une limite en a.*

Démonstration. Soient y et y' des limites de f en a, ainsi que des voisinages V et V' de y et y'. Nous prenons également les voisinages W et W' correspondants :

$$\begin{cases} f(W \cap A) \subset V & \text{(6.66a)} \\ f(W' \cap A) \subset V'. & \text{(6.66b)} \end{cases}$$

19. Définition 6.7.

Quitte à prendre des sous-ensembles nous pouvons supposer que W et W' sont ouverts. L'ensemble $W \cap W'$ est un ouvert contenant a et intersecte donc A. L'ensemble $(W \cap W') \cap A$ est non vide et

$$\begin{cases} f(W \cap W' \cap A) \subset f(W \cap A) \subset V & (6.67a) \\ f(W \cap W' \cap A) \subset f(W' \cap A) \subset V'. & (6.67b) \end{cases}$$

Donc les ensembles V et V' ont une intersection. Au final nous avons prouvé que deux voisinages de y et y' ont forcément une intersection ; étant donné que Y est séparé, nous devons avoir $y = y'$. □

La définition suivante est *la* définition de la continuité dans tous les cas.

Définition 6.69 (Fonction continue)**.**
*Une fonction $f : X \to Y$ est **continue au point** $a \in X$ si $f(a)$ est une limite*[20] *de f en a.*

*Une fonction $f : X \to Y$ est **continue** si pour tout ouvert $\mathcal{O}$ de Y, l'ensemble $f^{-1}(\mathcal{O})$ est ouvert dans X.*

La proposition 6.134 donnera des détails sur ce qu'il se passe lorsque l'espace est métrique.

Théorème 6.70.
Une fonction $f : X \to Y$ est une fonction continue si et seulement si elle est continue en chacun des points de X.

Démonstration. **Sens direct** Nous supposons que f est une fonction continue. Soit $a \in X$ et V, un voisinage de $f(a)$. Nous considérons $\mathcal{O}$, un voisinage ouvert de $f(a)$ contenu dans V ; l'ensemble $f^{-1}(\mathcal{O})$ est alors un ouvert contenant a, et l'image de $f^{-1}(\mathcal{O})$ par f est bien entendu contenue dans V.

Sens inverse Soit $\mathcal{O}$ un ouvert de Y. Pour prouver que $f^{-1}(\mathcal{O})$ est un ouvert de$X$, nous allons considérer un élément $a \in f^{-1}(\mathcal{O})$ et montrer qu'il existe un voisinage ouvert de a contenu dans $f^{-1}(\mathcal{O})$; le théorème 6.3 nous assurera alors que $f^{-1}(\mathcal{O})$ est ouvert.

L'ensemble $\mathcal{O}$ est un voisinage ouvert de $f(a)$ parce que a a été choisit dans $f^{-1}(\mathcal{O})$. Donc la continuité de f en a[21] nous assure qu'il existe un voisinage W de a tel que $f(W) \subset \mathcal{O}$. En prenant un ouvert contenant a à l'intérieur de W nous avons un voisinage ouvert de a contenu dans $f^{-1}(\mathcal{O})$.

□

Remarque 6.71.
À cause de l'éventuelle non unicité de la limite, deux fonctions continues et égales sur un sous-ensemble dense ne sont pas spécialement égales. Ce sera vrai sur les espaces métriques et plus généralement pour les espaces séparés. Voir la remarque 6.66.

6.7 Compacité

Quelques liens vers la notion de compacité.

(1) La définition d'un ensemble compact est la définition 6.10.

(2) Les compacts sont les fermés bornés par le théorème 6.123.

6.7.1 Quelque propriétés

Définition 6.72.
*Une famille $\mathcal{A}$ de parties de X a la **propriété d'intersection finie non vide** si tout sous-ensemble fini de $\mathcal{A}$ a une intersection non vide.*

Proposition 6.73.
Soit X un espace topologique et $K \subset X$. Les propriétés suivantes sont équivalentes :

20. Définition 6.65.
21. Définition 6.69.

(1) K est compact.

(2) Si $\{F_i\}$ est une famille de fermés telle que $K\bigcap_{i\in I}F_i=\varnothing$, alors il existe un sous-ensemble fini A de I tel que $K\bigcap_{i\in A}F_i=\varnothing$.

(3) Si $\{F_i\}_{i\in I}$ est une famille de fermés telle que $K\bigcap_{i\in A}F_i\neq\varnothing$ pour tout choix de A fini dans I, alors l'intersection complète est non vide : $K\bigcap_{i\in I}F_i\neq\varnothing$.

(4) Toute famille ayant la propriété d'intersection finie non vide a une intersection non vide.

Démonstration. Les propriétés (3) et (2) sont équivalentes par contraposition. De plus le point (4) est une simple reformulât ion en français de la propriété (3).

Prouvons (1) $\Rightarrow$ (2). Soit $\{F_i\}_{i\in I}$ une famille de fermés tels que $K\bigcap_{i\in I}F_i=\varnothing$. Les complémentaires $\mathcal{O}_i$ de F_i dans X recouvrent K et donc on peut en extraire un sous-recouvrement fini :

$$K\subset\bigcup_{i\in A}\mathcal{O}_i \tag{6.68}$$

pour un certain sous-ensemble fini A de I. Pour ce même choix A, nous avons alors aussi

$$K\bigcap_{i\in A}F_i=\varnothing. \tag{6.69}$$

L'implication (2) $\Rightarrow$ (1) est la même histoire. □

Proposition 6.74 ([69])**.**
Tout compact d'un espace topologique séparé est fermé.

Démonstration. Soit X un espace séparé et K compact dans X. Nous considérons $x\in\complement K$ et, par hypothèse de séparation, pour chaque $x\in K$ nous considérons un voisinage ouvert V_x de x et un voisinage ouvert W_x de y tels que $V_x\cap W_x=\varnothing$. Bien entendu les V_x forment un recouvrement ouvert de K dont nous pouvons extraire un sous-recouvrement fini : soit S fini dans K tel que

$$K\subset\bigcup_{x\in S}V_x. \tag{6.70}$$

L'ensemble $W=\bigcap_{x\in S}W_x$ est une intersection finie d'ouverts autour de y et est donc un ouvert autour de y. De plus W n'intersecte pas K parce que si $x\in K$ alors $x\in V_s$ pour un certain $s\in S$ par conséquent x n'est pas dans W_s et donc pas non plus dans l'intersection.

L'ouvert W prouve que y est dans l'intérieur du complémentaire de K, ce qui nous permet de conclure que le complémentaire de K est ouvert (théorème 6.3), aka que K est fermé. □

Lemme 6.75 ([70])**.**
Une partie fermée d'une compact est elle-même compacte.

Démonstration. Soit F fermé dans un compact K et $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de F par des ouverts. Vu que F est fermé, F^c est ouvert et $\{\mathcal{O}_i\}_{i\in I}\cup\{K\backslash F\}$ est un recouvrement de K par des ouverts. Si nous en extrayons un sous-recouvrement fini, c'est un recouvrement de F, et en supprimant éventuellement l'ouvert $K\backslash F$, ça reste un sous-recouvrement fini de F tout en étant extrait de $\{\mathcal{O}_i\}_{i\in I}$. □

Proposition 6.76.
Si V est une partie de l'espace topologique X muni de la topologie induite τ_V de celle de X, et si K est un compact de (V,τ_V) alors K est un compact de (X,τ_X).

Démonstration. Soient $(\mathcal{O}_\alpha)_{\alpha\in A}$ des ouverts de X recouvrant K. Alors les ensembles $V\cap\mathcal{O}_\alpha$ recouvrent également K, mais sont des ouverts de V. Donc il en existe un sous-recouvrement fini. Soient donc $(V\cap\mathcal{O}_i)_{i\in I}$ recouvrant K avec I un sous-ensemble fini de A. Les ensembles $(\mathcal{O}_i)_{i\in I}$ recouvrent encore K et sont des ouverts de X. □

Théorème 6.77.
L'image d'un compact par une fonction continue est un compact

Dans le cadre des espaces vectoriels normés, ce théorème est démontré en la proposition 6.129.

Démonstration. Soit $K \subset X$, un ensemble compact, et regardons $f(K)$; en particulier, nous considérons Ω, un recouvrement de $f(K)$ par des ouverts. Nous avons que

$$f(K) \subseteq \bigcup_{\mathcal{O}\in\Omega} \mathcal{O}. \tag{6.71}$$

Par construction, nous avons aussi

$$K \subseteq \bigcup_{\mathcal{O}\in\Omega} f^{-1}(\mathcal{O}), \tag{6.72}$$

en effet, si $x \in K$, alors $f(x)$ est dans un des ouverts de Ω, disons $f(x) \in \mathcal{O}_0$, et évidemment, $x \in f^{-1}(\mathcal{O})$. Les $f^{-1}(\mathcal{O})$ recouvrent le compact K, et donc on peut en choisir un sous-recouvrement fini, c'est à dire un choix de $\{f^{-1}(\mathcal{O}_1), \ldots, f^{-1}(\mathcal{O}_n)\}$ tels que

$$K \subseteq \bigcup_{i=1}^{n} f^{-1}(\mathcal{O}_i). \tag{6.73}$$

Dans ce cas, nous avons que

$$f(K) \subseteq \bigcup_{i=1}^{n} \mathcal{O}_i, \tag{6.74}$$

ce qui prouve la compacité de $f(K)$. □

6.8 Topologie induite

Définition 6.78.
Soit X un espace topologique et $A \subset X$. L'ensemble A devient un espace topologique en lui-même par la ***topologie induite*** *de X. Un ouvert de A est un ensemble de la forme $A \cap \mathcal{O}$ où $\mathcal{O}$ est un ouvert de X.*

Lemme 6.79.
Si $B \subset A$ alors la fermeture de B pour la topologie de A (induite de X) que nous noterons $\tilde{B}$ est

$$\tilde{B} = \bar{B} \cap A \tag{6.75}$$

où $\bar{B}$ est la fermeture de B pour la topologie de X.

Démonstration. Si $a \in \bar{B} \cap A$, un ouvert de A autour de a est un ensemble de la forme $\mathcal{O} \cap A$ où $\mathcal{O}$ est un ouvert de X. Vu que $a \in \bar{B}$, l'ensemble $\mathcal{O}$ intersecte B et donc $(\mathcal{O} \cap A) \cap B \neq \varnothing$. Donc a est bien dans l'adhérence de B au sens de la topologie de A.

Pour l'inclusion inverse, soit $a \in \tilde{B}$, et montrons que $a \in \bar{B} \cap A$. Par définition $a \in A$, parce que $\tilde{B}$ est une fermeture dans l'espace topologique A. Il faut donc seulement montrer que $a \in \bar{B}$. Soit donc $\mathcal{O}$ un ouvert de X contenant a ; par hypothèse $\mathcal{O} \cap A$ intersecte B (parce que tout ouvert de A contenant a intersecte B). Donc $\mathcal{O}$ intersecte B. Cela signifie que tout ouvert (de X) contenant a intersecte B, ou encore que $a \in \bar{B}$. □

Exemple 6.80
Si A est un ouvert de X, on pourrait croire que la topologie induite n'a rien de spécial. Il est vrai que B sera ouvert dans A si et seulement s'il est ouvert dans X, mais des choses se passent quand même. Prenons $X = \mathbb{R}$ et $A =]0, 1[$. Si $B =]\frac{1}{2}, 1[$, alors la fermeture de B dans A sera $\tilde{B} = [\frac{1}{2}, 1[$ et non $[\frac{1}{2}, 1]$ comme on l'aurait dans $\mathbb{R}$. △

Prendre la topologie induite de $\mathbb{R}$ vers un fermé de $\mathbb{R}$ donne des boules un peu spéciales comme le montre l'exemple suivant.

Exemple 6.81
Quid de la boule ouverte $B(1,\epsilon)$ dans le compact $[0,1]$? Par définition c'est

$$B(1,\epsilon) = \{x \in [0,1] \text{ tel que } |x-1| < \epsilon\} =]1-\epsilon, 1]. \tag{6.76}$$

Oui, cela est *ouvert* dans $[0,1]$. C'est d'ailleurs un des ouverts de la topologie induite de $\mathbb{R}$ sur $[0,1]$.

Donc pour la topologie de $[0,1]$, toutes les boules ouvertes $B(x,\delta)$ avec $x \in [0,1]$ sont incluses à $[0,1]$. △

Lemme 6.82.
Soit $A \subset X$ muni de la topologie induite de X et (x_n) une suite dans A. Si $x_n \xrightarrow{A} x$, alors $x_n \xrightarrow{X} x$.

Démonstration. Soit $\mathcal{O}$ un ouvert autour de x dans X. Alors $A \cap \mathcal{O}$ est un ouvert autour de x dans A et il existe $N \in \mathbb{N}$ tel que si $n \geqslant N$, alors $x_n \in A \cap \mathcal{O} \subset \mathcal{O}$. □

Lemme 6.83 ([1])**.**
Soit (X, τ_X) un espace topologique et $S \subset X$, un fermé de X sur lequel nous considérons la topologie induite τ_S. Si F est un fermé de (S, τ_S) alors F est fermé de (X, τ_X).

Démonstration. Nous savons que le complémentaire de F dans S est un ouvert de (S, τ_S) : il existe un ouvert $\Omega \in \tau_X$ tel que $S\backslash F = S \cap \Omega$. Si maintenant nous considérons le complémentaire de S dans X nous avons

$$F^c = (S\backslash F) \cup (X\backslash S) = (S \cap \Omega) \cup S^c = (S \cap \Omega) \cup (S^c \cap \Omega) \cup S^c = \Omega \cup S^c. \tag{6.77}$$

Vu que Ω et S^c sont des ouverts de X, l'union est un ouvert. Donc $F^c \in \tau_X$ et F est un fermé de X. □

6.9 Connexité

Dès qu'un ensemble est muni d'une métrique, nous pouvons définir les boules ouvertes, les voisinages et les sous-ensembles ouverts. Dès que l'on a identifié les sous-ensemble ouverts de E, nous disons que E devient un **espace topologique**. Nous allons maintenant un pas plus loin.

Définition 6.84.
*Lorsque E est un espace topologique, nous disons qu'un sous-ensemble A est **non connexe** quand on peut trouver des ouverts O_1 et O_2 disjoints tels que*

$$A = (A \cap O_1) \cup (A \cap O_2), \tag{6.78}$$

et tels que $A \cap O_1 \neq \varnothing$, et $A \cap O_2 \neq \varnothing$. Si un sous-ensemble n'est pas non-connexe, alors on dit qu'il est connexe.

Une autre façon d'exprimer la condition (6.78) est de dire que A n'est pas connexe quand il est contenu dans la réunion de deux ouverts disjoints qui intersectent tous les deux A.

Proposition 6.85.
Soit X un espace topologique. Les conditions suivantes sont équivalentes.

(1) L'espace X est connexe.

(2) Si $X = A \sqcup B$ avec A et B fermés dans X, alors $A = \varnothing$ ou $B = \varnothing$.

(3) Si $A \subset X$ avec A ouvert et fermé en même temps, alors $A = \varnothing$ ou $A = X$.

(4) Toute application continue $X \to \mathbb{Z}$ est constante.

Proposition 6.86.
L'image d'un ensemble connexe par une fonction continue est connexe.

Démonstration. Soit $f\colon X \to Y$ une application continue entre deux espaces topologiques, et E une partie connexe de X. Nous devons montrer que $f(E)$ est connexe dans Y.

Par l'absurde nous considérons A et B, deux ouverts de Y disjoints recouvrant $f(E)$. Étant donné que f est continue, les ensembles $f^{-1}(A)$ et $f^{-1}(B)$ sont ouverts dans X. De plus ces deux ensembles recouvrent E.

Si x est un élément de $f^{-1}(A) \cap f^{-1}(B)$, alors $f(x) \in A \cap B$, ce qui est impossible parce que nous avons supposé que A et B étaient disjoints. Par conséquent $f^{-1}(A)$ et $f^{-1}(B)$ sont deux ouverts disjoints recouvrant E. Contradiction avec la connexité de E. Nous concluons que $f(E)$ est connexe. □

Une application de ce théorème sera le théorème de valeurs intermédiaires 12.21.

Exemple 6.87
Les espaces topologiques $\mathbb{R}$ et $\mathbb{R}^2$ ne sont pas homéomorphes. △

Démonstration. Soit $f\colon \mathbb{R} \to \mathbb{R}^2$ un homéomorphisme. Nous posons $E = f\big(\mathbb{R}\backslash\{0\}\big)$ et $z_0 = f(0)$. Vu que f est bijective nous avons

$$E = \mathbb{R}^2\backslash\{z_0\}, \tag{6.79}$$

qui est connexe.

Vu que E est connexe et que f^{-1} est continue, la proposition 6.86 nous dit que $f^{-1}(E)$ est connexe. Mais par définition, $f^{-1}(E) = \mathbb{R}\backslash\{0\}$ qui n'est pas connexe. □

Proposition 6.88.
Si $A \subset X$ est connexe et si $A \subset B \subset \bar{A}$, alors B est connexe.

Proposition 6.89.
Stabilité de la connexité par union.

(1) Une union quelconque ce connexes ayant une intersection non vide est connexe.

(2) Si $A_1, \dots, A_n$ sont des connexes de X avec $A_i \cap A_{i+1} \neq \varnothing$, alors l'union $\bigcup_{n=1}^{n} A_i$ est connexe.

Démonstration. (1) Soient $\{C_i\}_{i\in I}$ un ensemble de connexes et un point p dans l'intersection : $p \in \bigcap_{i\in I} C_i$. Supposons que l'union ne soit pas connexe. Alors nous considérons A et B, deux ouverts disjoints recouvrant tous les C_i et ayant chacun une intersection non vide avec l'union.

Supposons pour fixer les idées que $p \in A$ et prenons $x \in B \cap \bigcup_{i\in I} C_i$. Il existe un $j \in I$ tel que $x \in C_j$. Avec tout cela nous avons

(a) $C_j \subset A \cup B$,

(b) $C_j \cap A \neq \varnothing$ parce que p est dans l'intersection,

(c) $C_j \cap B \neq \varnothing$ parce que x est dans cette intersection.

Cela contredit le fait que C_j soit connexe.

(2) Pour la seconde partie nous procédons de proche en proche. D'abord $A_1 \cup A_2$ est connexe par la première partie, ensuite $(A_1 \cup A_2) \cup A_3$ est connexe parce que les connexes $A_1 \cap A_2$ et A_3 ont un point d'intersection par hypothèse, et ainsi de suite.

□

6.10 Un peu de topologie réelle

Afin de pouvoir étudier la topologie des espaces métriques, il faut savoir quelques propriétés des réels parce que nous allons étudier la fonction distance qui est une fonction continue à valeurs dans les réels.

6.10.1 Suites numériques

La propriété suivante est le plus souvent donnée comme une définition lorsque seule la topologie sur $\mathbb{R}$ est considérée.

Proposition 6.90 (Limite d'une suite numérique)**.**
La suite (x_n) est convergente si et seulement s'il existe un réel ℓ tel que

$$\forall \epsilon > 0, \exists N \in \mathbb{N} \text{ tel que } \forall n \geqslant N, |x_n - \ell| < \epsilon. \tag{6.80}$$

Dans ce cas, le nombre ℓ est la limite de la suite (x_n). Nous dirons aussi souvent que la suite ***converge*** *vers le nombre ℓ.*

Démonstration. La limite d'une suite dans un espace topologique est la définition 6.4.

Si $x_n \to \ell$ et si $\epsilon > 0$ il existe N_ϵ tel que pour tout $n \geqslant N$ nous avons $x_n \in B(\ell, \epsilon)$ (parce que cette boule est un ouvert contenant ℓ). Vu la définition d'une boule (définition 6.19) et de la norme sur $\mathbb{R}$ (par 6.55), cette condition est bien $|x_n - \ell| < \epsilon$.

Dans l'autre sens, soit $\mathcal{O}$ un ouvert contenant ℓ. Par définition de la topologie, il existe $\epsilon > 0$ tel que $B(\ell, \epsilon) \subset \mathcal{O}$. La condition (6.80) nous assure qu'il existe N_ϵ tel que pour tout $n \geqslant N_\epsilon$ nous ayons

$$x_n \in B(\ell, \epsilon) \subset \mathcal{O}, \tag{6.81}$$

ce qui assure que la suite (x_n) converge vers ℓ pour la topologie métrique de $\mathbb{R}$. □

Une façon équivalente d'exprimer le critère (6.80) est de dire que pour tout ϵ positif, il existe un rang $N \in \mathbb{R}$ tel que l'intervalle $[\ell - \epsilon, \ell + \epsilon]$ contient tous les termes x_n au-delà de N.

Il est à noter que le rang N dont il est question dans la définition de suite convergente dépend de ϵ.

Exemple 6.91
Quelques suites usuelles.

(1) La suite $x_n = \frac{1}{n}$ converge vers 0.

(2) La suite $x_n = (-1)^n$ ne converge pas.

△

Une suite est dite **contenue** dans un ensemble A si $x_n \in A$ pour tout n. Une suite est **bornée supérieurement** s'il existe un M tel que $x_n \leqslant M$ pour tout n. De la même manière, la suite est bornée inférieurement s'il existe un m tel que $x_n \geqslant m$ pour tout n.

Le lemme suivant est souvent utilisé pour prouver qu'une suite est convergente.

Lemme 6.92.
Une suite croissante et bornée supérieurement converge. Une suite décroissante bornée inférieurement est convergente.

Théorème 6.93.
Toute suite réelle contenue dans un compact admet une sous-suite convergente.

Démonstration. Soit (x_n) une suite contenue dans la partie bornée $A \subset \mathbb{R}$. Nous disons qu'un élément x_k de la suite est *maximal* s'il est plus grand ou égal que tous les suivants : $x_k \geqslant x_{k'}$ dès que $k' \geqslant k$.

Si la suite a un nombre infini d'éléments maximaux, alors la suite des éléments maximaux est décroissante. Si nous n'avons qu'un nombre fini de points maximaux, alors la suite de départ est croissante à partir du dernier point maximal.

Dans les deux cas nous avons trouvé une sous-suite des x_n qui est monotone (décroissante ou croissante selon le cas), et donc convergente parce que contenue dans un borné (lemme 6.92). □

Lemme 6.94.
Pour tout $\alpha > 0$ et $a < 1$ nous avons la limite

$$\lim_{n\to\infty} n^\alpha a^n = 0 \tag{6.82}$$

Démonstration. Soit $k \in \mathbb{N}$ plus grand que α. Soit la suite numérique $s_n = n^k a^n$. Tous ses termes sont positifs et

$$\frac{s_n}{s_{n+1}} = \left(\frac{n}{n+1}\right)^k \frac{1}{a}. \tag{6.83}$$

Étant donne que $n/n+1 \to 1$ et que $a < 1$, il existe un certain rang à partir duquel la suite (s_n) est décroissante. Deux conclusions :

— Elle est majorée par une constante M.
— Elle est convergente par le lemme 6.92.

Soit l tel que $ka^l < 1$ et $n > l$ alors

$$s_{n+l} = (n+l)^k a^{n+l} \leqslant k n^k a^n a^l = k a^l s_n \leqslant k a^l M. \tag{6.84}$$

La majoration est due au fait que dans $(n+l)^k$ nous avons k termes tous plus petits que n^k. De la même façon,

$$s_{2n+2l} \leqslant k a^{2l} s_{2n} \leqslant k a^{2l} M. \tag{6.85}$$

En posant $\varphi(i) = in + il$ nous avons

$$s_{\varphi(i)} \leqslant k a^i M, \tag{6.86}$$

qui est une sous-suite convergente vers 0. Or si une suite est convergente (ce qui est le cas de (s_n)), toutes les sous-suites convergent vers la même limite. Nous en concluons que $s_n \to 0$. □

Remarque 6.95.
Nous ne pouvons pas prouver cette limite en utilisant les propriétés de l'exponentielle ou du logarithme (en particulier les dérivées comme on voudrait en utiliser pour la règle de l'Hospital ou pour un développement limité) parce que les exponentielles sont définies par leur série entière et que leurs propriétés de dérivation dépendent de la proposition 15.27.

Proposition 6.96.
Une suite (x_n) dans $\mathbb{R}^m$ est convergente dans $\mathbb{R}^m$ si et seulement si les suites de chaque composantes sont convergentes dans $\mathbb{R}$. Dans ce cas nous avons

$$\lim x_n = \Big(\lim(x_n)_1, \lim(x_n)_2, \dots, \lim(x_n)_m\Big) \tag{6.87}$$

où $(x_n)_k$ dénote la k-ième composante de (x_n).

Exemple 6.97
La suite $x_n = \left(\frac{1}{n}, 1 - \frac{1}{n}\right)$ converge vers $(0,1)$ dans $\mathbb{R}^2$. En effet, en utilisant la proposition 6.96, nous devons calculer séparément les limites

$$\begin{aligned} \lim \frac{1}{n} &= 0 \\ \lim \left(1 - \frac{1}{n}\right) &= 1. \end{aligned} \tag{6.88}$$

△

Exemple 6.98
Étant donné que la suite $(-1)^n$ n'est pas convergente, la suite $x_n = \left((-1)^n, \frac{1}{n}\right)$ n'est pas convergente dans $\mathbb{R}^2$. △

6.10.2 Maximum, supremum et compagnie

Ce n'est un secret pour personne que $\mathbb{R}$ est un ensemble totalement ordonné[22] : il y a des éléments plus grands que d'autres, et mieux : à chaque fois que je prends deux éléments différents dans $\mathbb{R}$, il y en a un des deux qui est plus grand que l'autre. Il n'y a pas d'*ex æquo* dans $\mathbb{R}$.

Si je regarde l'intervalle $I = [0, 1]$, je peux même dire que 10 est plus grand que tous les éléments de I. Nous disons que 10 est un *majorant* de $[0, 1]$. La définition est la suivante.

Définition 6.99.
*Lorsque A est un sous-ensemble de $\mathbb{R}$, on dit que s est un **majorant** de A si s est plus grand que tous les éléments de A. En d'autres termes, si*

$$\forall x \in A,\ s \geqslant x.$$

Je me permet d'insister sur le fait que l'inégalité n'est pas stricte. Ainsi, 1 est un majorant de $[0, 1]$. Dès qu'un ensemble a un majorant, il en a plein. Si s majore l'ensemble A, alors évident $s + 1$, $s + 4$, $s + \pi^2$ majorent également A.

Exemple 6.100
Une petite galerie d'exemples de majorants.
- L'intervalle fermé $[4, 8]$ admet entre autres 8 et 130 comme majorants,
- l'intervalle ouvert $]4, 8[$ admet également 8 et 130 comme majorants,
- 7 n'est pas un majorant de $[1, 5]\cup]8, 32]$,
- 10/10 majore les notes qu'on peut obtenir à un devoir.
- l'intervalle $[4, \infty[$ n'a pas de majorants.

△

Définition 6.101.
*Le **supremum** d'un ensemble est le plus petit majorant. En d'autres terme, s est un supremum de A si tout nombre plus petit que s ne majore pas A, ou encore,*

$$\forall x < s, \exists y \in A \text{ tel que } y > x.$$

*Nous disons que M est un **maximum** de A si M est un supremum* et *$M \in A$.*

Quand s est un supremum de A, ça veut dire que le moindre pas vers la gauche que l'on fait à partir de s (c'est à dire le moindre ϵ), et on tombe dans A, ou tout au moins, il existe des éléments de A qui sont plus grand que $s - \epsilon$.

6.10.2.1 . . . et quelque exemples

En matières de notations, le maximum de l'ensemble A est noté $\max A$, le supremum est noté $\sup A$. Le minimum et l'infimum sont notés $\min A$ et $\inf A$.

Exemple 6.102
Exemples de différence entre majorant, supremum et maximum.
- Le nombre 10 est un supremum, majorant et maximum de l'intervalle fermé $[0, 10]$,

22. Définition 2.45.

— Le nombre 10 est un majorant et un supremum, mais pas un maximum de l'intervalle ouvert $]0, 10[$,
— Le nombre 136 est un majorant, mais ni un maximum ni un supremum de l'intervalle $[0, 10]$.

△

En utilisant les notations concises, ces différents cas s'écrivent ainsi :

$$10 = \max[0, 10] = \sup[0, 10] \qquad 10 = \sup[0, 10[$$

Exemple 6.103
Si on dit que un pont s'effondre à partir d'une charge de 10 tonnes, alors 10 tonnes est un *supremum* des charges que le pont peut supporter : si on met $9,999999$ tonnes dessus, il tient encore le coup, mais si on ajoute un gramme, alors il s'effondre (on sort de l'ensemble des charges acceptables).
△

Exemple 6.104
Si on dit qu'un pont résiste jusqu'à 10 tonnes, alors 10 tonnes est un *maximum* de la charge acceptable. Sur ce pont-ci, on peut ajouter le dernier gramme. Mais à partir de là, le moindre truc qu'on ajoute, il s'effondre. △

Lorsque vous lisez que la charge maximale d'un camion est de 2.5 tonnes, est-ce que cela veut dire que vous pouvez y mettre 2.5 tonnes, mais qui si un oiseau se pose dessus, le camion s'effondre ? Ou bien est-ce que cela signifie qu'à 2.5 tonnes le camion s'écroule, mais que toute charge inférieure est valable ?

C'est à cette rude question que nous allons nous attaquer maintenant.

Définition 6.105.
Soit une partie A de $\mathbb{R}$. Nous disons qu'un nombre M est un ***majorant*** *de A si M est plus grand ou égal que tous les éléments de A, c'est à dire si*

$$\forall a \in A,\ M \geqslant a. \tag{6.89}$$

Un ***minorant*** *de A est un nombre m tel que*

$$\forall a \in A,\ m \leqslant a. \tag{6.90}$$

Définition 6.106.
Soit A une partie majorée de $\mathbb{R}$. Le ***supremum*** *de A est le plus petit des majorants, c'est à dire le nombre M tel que*

(1) $M \geqslant x$ pour tout $x \in A$,
(2) pour tout ε, le nombre $M - \varepsilon$ n'est pas un majorant de a, c'est à dire qu'il existe un élément $x \in A$ tel que $x > M - \varepsilon$.

Nous notons $\sup A$ *le supremum de A.*

De la même façon, ***l'infimum*** *de A, noté* $\inf A$*, est le plus grand de ses minorants.*

Par convention, si la partie n'est pas bornée vers le haut, nous dirons que son supremum n'existe pas, ou bien qu'il est égal à $+\infty$, suivant les contextes. Pour votre culture générale, sachez toutefois que $\infty \notin \mathbb{R}$.

La définition est justifiée par le lemme 6.107 et la proposition 6.109. Le premier montre que si A possède un infimum, alors il est unique, tandis que le second montre que toute partie majorée de $\mathbb{R}$ accepter un supremum, et que toute partie minorée accepte un infimum.

Lemme 6.107.
Soit A une partie de $\mathbb{R}$. Supposons que m_1 et m_2 soient deux nombres qui vérifient les propriétés de l'infimum de A. Alors $m_1 = m_2$.

Démonstration. Si $_1 \neq m_2$, nous pouvons supposer $m_2 > m_1$. Dans ce cas, étant donné que m_1 est un infimum, m_2 ne peut pas minorer A, et donc ne peut pas être un infimum. □

Maintenant il est important de se rendre compte d'une chose : un ensemble ne peut avoir qu'un seul maximum et supremum. Jusqu'à présent nous avons toujours dit *un* supremum. À partir de maintenant nous pouvons dire *le* supremum. La preuve de cela est assez simple.

Proposition 6.108.
Si A est un sous-ensemble de $\mathbb{R}$ admettant un supremum, alors il n'a qu'un seul supremum ; et s'il accepte un maximum, il n'en accepte un seul, et le maximum est égal au supremum.

Démonstration. Commençons par l'affirmation concernant le supremum. Supposons que x et y soient tous les deux suprema différents de A. Étant donné que $x \neq y$, nous pouvons supposer que $x < y$, et donc, par définition du fait que y est un supremum, il existe un élément de A qui est plus grand que x. Cela contredit le fait que x soit supremum. En conclusion, il ne peut pas y avoir deux suprema différents pour un même ensemble.

Étant donné qu'un maximum est un supremum, il ne peut pas y avoir deux maxima différents vu qu'il ne peut pas y avoir deux suprema différents. □

Proposition 6.109.
Tout sous-ensemble de $\mathbb{R}$ borné vers le bas possède un infimum ; tout sous-ensemble de $\mathbb{R}$ borné vers le haut possède un supremum.

La preuve qui suit est proche de celle donnée par Wikipédia dans l'article Least uppert bound principle.

Démonstration. Soit A, une partie de $\mathbb{R}$. Nous allons trouver son infimum en suivant une méthode de dichotomie. Pour cela nous allons construire trois suites en même temps de la façon suivante. D'abord nous choisissons un point x_0 de A et un point x_1 qui minore A (qui existe par hypothèse) :

$$\begin{aligned} &x_0 \text{ est un élément de } A,\\ &x_1 \text{ est un minorant de } A,\\ &a_0 = x_0\\ &b_0 = x_1\\ &b_1 = x_1. \end{aligned} \tag{6.91}$$

Ensuite, nous faisons la récurrence suivante :

$$\begin{aligned} x_{n+1} &= \frac{a_n + b_n}{2},\\ a_{n+1} &= \begin{cases} a_n & \text{si } x_{n+1} \text{ minore } A\\ x_{n+1} & \text{sinon,} \end{cases}\\ b_{n+1} &= \begin{cases} x_{n+1} & \text{si } x_{n+1} \text{ minore } A\\ b_n & \text{sinon.} \end{cases} \end{aligned} \tag{6.92}$$

Nous allons montrer que a_n et (b_n) sont des suites convergentes de même limite et que cette limite est l'infimum de A.

Soit $n \in \mathbb{N}$; il y a deux possibilités. Soit $a_n = a_{n-1}$ et $b_n = x_n$, soit $a_n = x_n$ et $b_n = b_{n-1}$. Supposons que nous soyons dans le premier cas (le second se traite de façon similaire). Alors nous avons

$$\begin{aligned} |a_n - b_n| &= |a_{n-1} - x_n|\\ &= \left|a_{n-1} - \frac{a_{n-1} + b_{n-1}}{2}\right|\\ &= \frac{1}{2}|a_{n-1} - b_{n-1}|, \end{aligned} \tag{6.93}$$

ce qui prouve que $|a_n - b_n| \to 0$. Nous montrons maintenant que la suite (a_n) est de Cauchy. En effet nous avons

$$|a_n - a_{n-1}| = \begin{cases} 0 \\ |\frac{a_n - b_n}{2}| \end{cases} \leqslant \frac{1}{2n}. \tag{6.94}$$

Il en est de même pour la suite (b_n). Ce sont deux suites de Cauchy (donc convergentes par le théorème 6.57) qui convergent vers la même limite. Soit ℓ cette limite.

Le nombre ℓ minore A. En effet si $a \in A$ est plus petit que ℓ, les éléments b_n tels que $|b_n - \ell| < |a - \ell|$ ne peuvent pas minorer A. D'autre part, pour tout ϵ, le nombre $\ell + \epsilon$ ne peut pas minorer A. En effet, ℓ est la limite de la suite décroissante (a_n), donc il existe a_n entre ℓ et $\ell + \epsilon$. Mais a_n ne minore pas A, donc $\ell + \epsilon$ ne minore pas non plus A.

Nous avons prouvé que toute partie minorée de $\mathbb{R}$ possède un infimum. La preuve que toute partie majorée possède un supremum se fait de la même façon.

□

Définition 6.110.
Si le supremum d'un ensemble appartient à l'ensemble, nous l'appelons ***maximum****. De la même façon si l'infimum d'un ensemble appartient à l'ensemble, nous disons que c'est le* ***minimum****.*

Exemple 6.111
Pour les intervalles, ces notions sont simples : les bornes de l'intervalle sont les supremum et infimum, et ce sont des minima et maxima si l'intervalle est fermé. Le nombre 53 est un majorant.

(1) $A = [1, 2]$. Tous les nombres plus petits ou égaux à 1 sont minorants, 1 est infimum et minimum. Le nombre 2 est un majorant, le maximum et le supremum.
(2) $B =]3, \pi[$. Le nombre π est le supremum et est un majorant, mais n'est pas le maximum (parce que $\pi \notin B$). L'ensemble B n'a pas de maximum. Bien entendu, -1000 est un minorant.

△

Il existe évidement de nombreux exemples plus vicieux.

Exemple 6.112
Prenons $E = \{\frac{1}{n} \text{ tel que } n \in \mathbb{N}_0\}$, dont les premiers points sont indiqués sur la figure 6.2. Cet ensemble est constitué des nombres 1, $\frac{1}{2}$, $\frac{1}{3}$, ... Le plus grand d'entre eux est 1 parce que tous les nombres de la forme $\frac{1}{n}$ avec $n \geqslant 1$ sont plus petits ou égaux à 1. Le nombre 1 est donc maximum de E.

L'ensemble E n'a par contre pas de minimum parce que tout élément de E s'écrit $\frac{1}{n}$ pour un certain n et est plus grand que $\frac{1}{n+1}$ qui est également dans E.

Prouvons que zéro est l'infimum de E. D'abord, tous les éléments de E sont strictement positifs, donc zéro est certainement un minorant de E. Ensuite, nous savons que pour tout $\varepsilon > 0$, il existe un n tel que $\frac{1}{n}$ est plus petit que ε. L'ensemble E possède donc un élément plus petit que $0 + \varepsilon$, et zéro est bien l'infimum. △

L'exemple suivant est une source classique d'erreurs en ce qui concerne l'infimum. Il sera à relire après avoir vu la définition de limite (définition 6.90).

Exemple 6.113
Les premiers points de l'ensemble $F = \{\frac{(-1)^n}{n} \text{ tel que } n \in \mathbb{N}_0\}$ sont représentés à la figure 6.3. Bien que (comme nous le verrons plus tard) la limite de la suite $x_n = (-1)^n/n$ soit zéro, il n'est pas correct de dire que zéro est l'infimum de l'ensemble F. Le dessin, au contraire, montre bien que -1 est le minium (aucun point est plus bas que -1), tandis que le maximum est $1/2$.

Nous reviendrons avec cet exemple dans la suite. Pour l'instant, ayez bien en tête que zéro n'est rien de spécial pour l'ensemble F en ce qui concerne les notions de maximum, minimum et compagnie. △

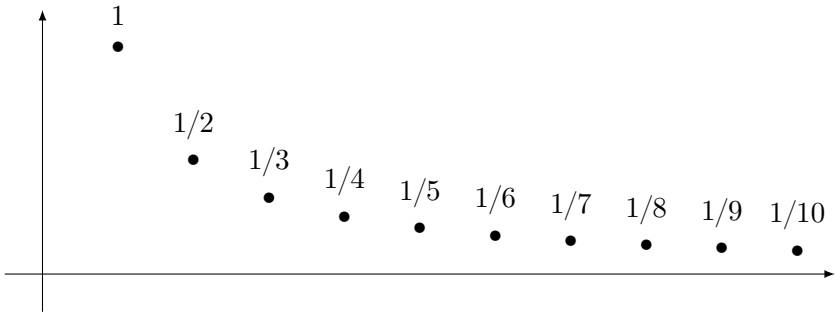


FIGURE 6.2 – Les premiers points du type $x_n = 1/n$.

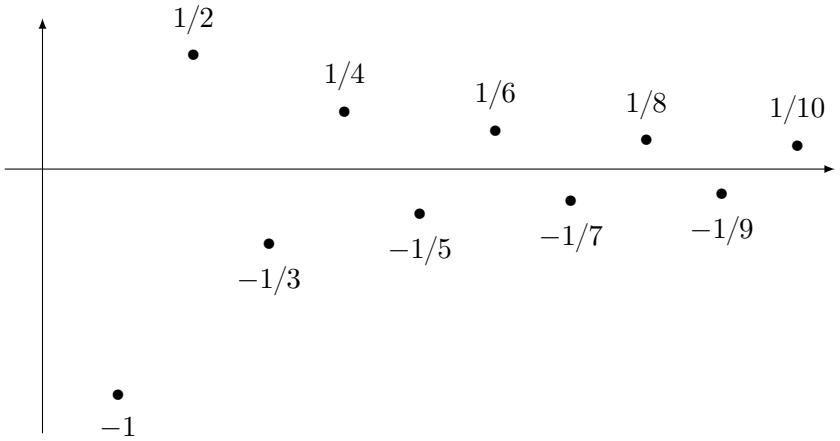


FIGURE 6.3 – Les quelques premiers points du type $(-1)^n/n$.

Lorsque $x \in E$, nous disons qu'un **voisinage** de x est n'importe quel sous-ensemble de E qui contient une boule ouverte centrée en x. Nous disons qu'un ensemble est **ouvert** s'il contient un voisinage de chacun de ses points. Par convention, nous disons que l'ensemble vide est ouvert.

Définition 6.114.
L'ensemble des boules ouvertes d'un espace métrique forment la ***topologie*** *de l'espace.*

Nous allons dire qu'une partie A d'un espace métrique est **bornée** s'il existe une boule [23] qui contient A.

Lemme 6.115.
Le supremum d'un ensemble ouvert n'est pas dans l'ensemble (et n'est donc pas un maximum).

Démonstration. Soit $\mathcal{O}$, un ensemble ouvert et s, son supremum. Si s était dans $\mathcal{O}$, on aurait un voisinage $B = B(s, r)$ de s contenu dans $\mathcal{O}$. Le point $s + r/2$ est alors à la fois dans $\mathcal{O}$ et plus grand que s, ce qui contredit le fait que s soit un supremum de $\mathcal{O}$. □

Par le même genre de raisonnements, on montre que l'union et l'intersection de deux ouverts sont encore des ouverts.

Remarque 6.116.
L'intersection d'une *infinité* d'ouverts n'est pas spécialement un ouvert comme le montre l'exemple suivant :

$$\mathcal{O}_i =]1, 2 + \frac{1}{i}[.$$

Tous les ensembles $\mathcal{O}_i$ contiennent le point 2 qui est donc dans l'intersection. Mais quel que soit

23. À titre d'exercice, je te laisse te convaincre que l'on peut dire boule *ouverte* ou *fermée* au choix sans changer la définition.

le $\epsilon > 0$ que l'on choisisse, le point $2 + \epsilon$ n'est pas dans $\mathcal{O}_{(1/\epsilon)+1}$. Donc aucun point au-delà de 2 n'est dans l'intersection, ce qui prouve que 2 ne possède pas de voisinages contenus dans $\cap_{i=1}^{\infty}\mathcal{O}_i$.

Proposition 6.117.
Prouver que, quels que soient les ensembles A et B dans $\mathbb{R}$, nous avons

$$\sup(A \cap B) \leqslant \sup A \leqslant \sup(A \cup B).$$

6.10.3 Intervalles

6.10.3.1 Connexité

Nous allons déterminer tous les sous-ensembles connexes[24] de $\mathbb{R}$. Pour cela nous relisons d'abord la notion d'intervalle donnée en 2.8. La partie $I \subset \mathbb{R}$ est un intervalle si pour tout $a, b \in I$, tout nombre entre a et b est également dans I. Cette définition englobe tous les exemples connus d'intervalles ouverts, fermés avec ou sans infini : $[a, b]$, $[a, b[$, $]-\infty, a]$, ... L'ensemble $\mathbb{R}$ lui-même est un intervalle.

Si I est un intervalle, les nombres $\inf(I)$ et $\sup(I)$[25] sont les **extrémités** de I.

Définition 6.118.
*Étant donnés deux points a et b dans $\mathbb{R}^p$ on appelle **segment** d'extrémités a et b, et on note $[a, b]$, l'image de $[0, 1]$ par l'application $s : [0, 1] \to \mathbb{R}^p$, $s(t) = (1 - t)a + tb$. On pose $]a, b[= s(]0, 1[)$, et $]a, b] = s(]0, 1])$.*

Il faut observer que le segment $[a, b]$ est une courbe orientée : certes en tant que ensembles, $[a, b] = [b, a]$, mais si nous regardons la fonction de t correspondante à $[b, a]$, nous voyons qu'elle va dans le sens inverse de celle qui correspond à $[a, b]$. Nous approfondirons ces questions lorsque nous parlerons d'arcs paramétrés autour de la section 19.7.

Le segment $[b, a]$ est l'image de l'application $r\colon [0, 1] \to \mathbb{R}^p$ donnée par $r(t) = (1 - t)b + ta$.

Une des nombreuses propositions qui vont servir à prouver le théorème des valeurs intermédiaires 12.21 est la suivante.

Proposition 6.119.
Une partie de $\mathbb{R}$ est connexe si et seulement si c'est un intervalle.

Démonstration. La preuve est en deux partie. D'abord nous démontrons que si un sous-ensemble de $\mathbb{R}$ est connexe, alors c'est un intervalle ; et ensuite nous démontrons que tout intervalle est connexe.

Afin de prouver qu'un ensemble connexe est toujours un intervalle, nous allons prouver que si un ensemble n'est pas un intervalle, alors il n'est pas connexe. Prenons A, une partie de $\mathbb{R}$ qui n'est pas un intervalle. Il existe donc a, $b \in A$ et un x_0 entre a et b qui n'est pas dans A. Comme le but est de prouver que A n'est pas connexe, il faut couper A en deux ouverts disjoints. L'élément x_0 qui n'est pas dans A est le bon candidat pour effectuer cette coupure. Prenons M, un majorant de A et m, un minorant de A, et définissons

$$\begin{aligned}\mathcal{O}_1 &=]m, x_0[\\ \mathcal{O}_2 &=]x_0, M[.\end{aligned}$$

Si A n'a pas de minorant, nous remplaçons la définition de $\mathcal{O}_1$ par $]-\infty, x_0[$, et si A n'a pas de majorant, nous remplaçons la définition de $\mathcal{O}_2$ par $]x_0, \infty[$. Dans tous les cas, ce sont deux ensembles ouverts dont l'union recouvre tout A. En effet, $\mathcal{O}_1 \cup \mathcal{O}_2$ contient tous les nombres entre un minorant de A et un majorant sauf x_0, mais on sait que x_0 n'est pas dans A. Cela prouve que A n'est pas connexe.

24. Définition 6.84.
25. Qui existent par la proposition 6.109, quitte à poser $\pm\infty$ comme infimum et supremum lorsque I n'est pas borné.

Jusqu'à présent nous avons prouvé que si un ensemble n'est pas un intervalle, alors il ne peut pas être connexe. Pour remettre les choses à l'endroit, prenons un ensemble connexe, et demandons-nous s'il peut être autre chose qu'un intervalle ? La réponse est *non* parce que s'il était autre chose, il ne serait pas connexe.

Prouvons à présent que tout intervalle est connexe. Pour cela, nous refaisons le coup de la contraposée. Nous allons donc prendre une partie A de $\mathbb{R}$, supposer qu'elle n'est pas connexe et puis prouver qu'elle n'est alors pas un intervalle. Nous avons deux ouverts disjoints $\mathcal{O}_1$ et $\mathcal{O}_2$ tels que $A \subset \mathcal{O}_1 \cup \mathcal{O}_2$. Prenons $a \in A_1$ et $b \in A_2$. Pour fixer les idées, on suppose que $a < b$. Maintenant, le jeu est de montrer qu'il existe une point x_0 entre a et b qui ne soit pas dans A (cela montrerait que A n'est pas un intervalle). Nous allons prouver que c'est le cas du point

$$x_0 = \sup\{x \in \mathcal{O}_1 \text{ tel que } x < b\}.$$

Étant donné que l'ensemble $\mathcal{A} = \{x \in \mathcal{O}_1 \text{ tel que } x < b\}$ est ouvert [26], le point x_0 n'est pas dans l'ensemble par le lemme 6.115. Nous avons donc

— soit x_0 n'est pas dans $\mathcal{O}_1$,
— soit $x_0 \leqslant b$,
— soit les deux en même temps.

Nous allons montrer qu'un tel x_0 ne peut pas être dans A. D'abord, remarquons que $\sup \mathcal{A} \leqslant \sup \mathcal{O}$ parce que $\mathcal{A}$ est une intersection de $\mathcal{O}$ avec quelque chose. Ensuite, il n'est pas possible que x_0 soit dans $\mathcal{O}_2$ parce que tout élément de $\mathcal{O}_2$ possède un voisinage contenu dans $\mathcal{O}_2$. Un point de $\mathcal{O}_2$ est donc toujours strictement plus grand que le supremum de $\mathcal{O}_1$.

Maintenant, remarque que si $x_0 \leqslant b$, alors $x_0 = b$, sinon b serait un majorant de $\mathcal{A}$ plus petit que x_0, ce qui n'est pas possible vu que x_0 est le supremum de $\mathcal{A}$ et donc le plus petit majorant. Oui mais si $x_0 = b$, c'est que $x_0 \in \mathcal{O}_2$, ce qu'on vient de montrer être impossible. Nous voila déjà débarrassé des deuxièmes et troisièmes possibilités.

Si la première possibilité est vraie, alors x_0 n'est pas dans A parce qu'on a aussi prouvé que $x_0 \notin \mathcal{O}_2$. Or n'être ni dans $\mathcal{O}_1$ ni dans $\mathcal{O}_2$ implique de ne pas être dans A. Ce point $x_0 = \sup \mathcal{A}$ est donc hors de A.

Oui, mais comme $a \in \mathcal{A}$, on a obligatoirement que $x_0 \geqslant a$. Mais par construction, on a aussi que $x_0 \leqslant b$ (ici, l'inégalité est même stricte, mais ce n'est pas important). Donc

$$a \leqslant x_0 \leqslant b$$

avec a, $b \in A$, et $x_0 \notin A$. Cela finit de prouver que A n'est pas un intervalle. □

Exemple 6.120
L'application

$$\begin{aligned} f\colon]0, 2\pi[&\to S^1 \\ x &\mapsto \begin{pmatrix} \cos(x) \\ \sin(x) \end{pmatrix} \end{aligned} \tag{6.95}$$

est un homéomorphisme. Vu que $]0, 2\pi[$ est connexe (proposition 6.119) la proposition 6.86 implique que le cercle privé d'un point est connexe. △

6.10.3.2 Compacité

Lemme 6.121 ([71])**.**
Si $a < b \in \mathbb{R}$ alors le segment $[a, b]$ est compact [27].

26. C'est l'intersection entre l'ouvert $\mathcal{O}_1$ et l'ouvert $\{x \text{ tel que } x < b\}$.
27. Définition 6.10

Démonstration. Soit $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de $[a,b]$ par des ouverts. Nous posons

$$M = \{x \in [a,b] \text{ tel que } [a,x] \text{ admet un sous-recouvrement fini extrait de } \{\mathcal{O}_i\}_{i\in I}\}. \tag{6.96}$$

Notre but est de prouver que $b \in M$.

a est dans M Le point a est naturellement dans un des $\mathcal{O}_i$. L'intervalle $[a,a]$ est donc recouvert par un seul des $\mathcal{O}_i$.

M est un intervalle Soit $m \in M$ et $m' \in [a,m[$. Le sous-recouvrement fini qui recouvre $[a,m]$ recouvre a fortiori $[a,m']$.

Les trois possibilités restantes À ce niveau de la preuve, il reste trois possibilités pour M soit il est de la forme $[a,c]$ ou $[a,c[$ avec $c<b$, soit il est de la forme $[a,b]$. Nous allons maintenant éliminer les deux premiers cas.

Ce que M n'est pas D'abord M n'est pas de la forme $[a,c[$ avec $c<b$. En effet nous commençons par considérer $\mathcal{O}_{i_0}$ un ouvert du recouvrement contenant c et $m<c$ dans $\mathcal{O}_{i_0}$. Si nous joignons $\mathcal{O}_{i_0}$ à un recouvrement fini de $[a,m]$ alors nous avons un recouvrement fini de $[a,c]$, et donc $c \in M$.

Ensuite M n'est pas de la forme $[a,c]$ avec $c<b$. En effet si on a un recouvrement fini de $[a,c]$ par des ouverts, alors un de ces ouverts contient c et donc contient des éléments de $[a,b]$ plus grands que c.

Nous déduisons que $M=[a,b]$ et qu'il est possible d'extraire un sous-recouvrement fini recouvrant $[a,b]$. □

Lemme 6.122 ([1])**.**
Si K_1 et K_2 sont des compacts dans $\mathbb{R}$ alors $K_1 \times K_2$ est compact dans $\mathbb{R}^2$.

Démonstration. Soit $\{\mathcal{O}_i\}_{i\in I}$ un recouvrement de $K_1 \times K_2$ par des ouverts ; grâce au lemme 6.63 nous pouvons supposer que ce sont des carrés. Pour chaque $x \in K_1$, l'ensemble $\{x\}\times K_2$ est compact et donc recouvert par un nombre fini des $\mathcal{O}_i$. Soit R_x un ensemble fini des $\mathcal{O}_i$ recouvrant $\{x\}\times K_2$.

Vu que R_x est une collection finie de carrés nous pouvons considérer m_x, le minimum des rayons. L'ensemble K_1 est recouvert par les boules $B(x,m_x)$ et il existe donc une collection finie de $\{x_i\}_{i\in A}$ tels que $B(x_i,m_{x_i})$ recouvre K_1.

Alors $\{R_{x_i}\}_{i\in A}$ recouvre $K_1\times K_2$ parce que R_{x_i} recouvre l'ensemble $B(x_i,m_{x_i})\times\{K_2\}$. □

Théorème 6.123 (Théorème de Borel-Lebesgue)**.**
Une partie d'un espace vectoriel normé réel de dimension finie est compacte si et seulement si elle est fermée et bornée.

Démonstration. Sens direct.

Compact implique borné En effet si K est non borné dans E alors K contient une suite (x_n) avec $\|x_n\|>n$. Les boules $B_i(x_i,\frac{1}{3})$ sont disjointes. On pose $\mathcal{O}_0 = \complement\bigcup_i \overline{B(x_i,\frac{1}{5})}$, qui est ouvert comme complément d'un fermé. Pour $i \geqslant 1$ nous posons $\mathcal{O}_i = B(x_i,\frac{1}{4})$. Nous avons

$$K \subset \bigcup_{i\in\mathbb{N}} \mathcal{O}_i \tag{6.97}$$

mais vu que x_i est uniquement dans $\mathcal{O}_i$, nous ne pouvons pas extraire de sous-recouvrement fini.

Compact implique fermé Cela est la proposition 6.74.

Sens réciproque.

Un intervalle fermé et borné est compact dans $\mathbb{R}$ C'est le lemme 6.121.

Un produit de segments est compact Le produit de deux compacts de $\mathbb{R}$ est un compact dans $\mathbb{R}^2$ par le lemme 6.122.

Un fermé et borné et compact Soit K fermé et borné. Vu que K est borné, il est contenu dans un produit de segments. L'ensemble K est donc compact parce que fermé dans un compact, lemme 6.75.

□

Exemple 6.124(Un compact non fermé)
En général, un compact n'est pas toujours fermé. Si nous prenons par exemple un ensemble X de plus de deux points muni de la topologie grossière $\{\emptyset, X\}$. Toutes les parties de cet espace sont compactes, mais les seuls fermés sont $\{\emptyset, X\}$. Toutes les autres parties sont alors compactes et non fermées. △

Exemple 6.125(Compacité de la boule unité)
La boule unité fermée $\overline{B(0,1)}$ d'un espace vectoriel normé de dimension finie est compacte parce que fermée et bornée. En dimension infinie, cela n'est plus le cas. Certes la boule unité est encore fermée et bornée, mais elle n'est plus compacte. En effet nous allons donner un recouvrement par des ouverts duquel il ne sera pas possible d'extraire un sous-recouvrement fini.

Autour de chacune des extrémités des vecteurs de base, nous considérons la boule $A_i = B(e_i, \frac{1}{3})$. Ensuite aussi l'ouvert

$$B(0,1)\backslash\bigcup_i \overline{B(e_i, \frac{1}{4})}. \tag{6.98}$$

Le tout recouvre $B(0,1)$ mais toutes les premières boules sont nécessaires. △ Le théorème de

Bolzano-Weierstrass nous permettra de prouver plus simplement la non compacité en dimension infinie. Voir l'exemple 6.148.

6.10.4 Image d'un compact

Un résultat important dans la théorie des fonctions sur les espaces vectoriels normés est qu'une fonction continue sur un compact est bornée et atteint ses bornes. Ce résultat sera (dans d'autres cours) énormément utilisé pour trouver des maxima et minima de fonctions. Le théorème exact est le suivant.

Théorème 6.126.
Une fonction à valeurs réelles continue sur un compact est bornée et atteint ses bornes.

C'est à dire qu'il existe $x_0 \in K$ tel que $f(x_0) = \inf\{f(x) \text{ tel que } x \in K\}$ ainsi que x_1 tel que $f(x_1) = \sup\{f(x) \text{ tel que } x \in K\}$.

Démonstration. Soit K un espace topologique compact et $f\colon K \to \mathbb{R}$ une fonction continue. Alors le théorème 6.77 indique que $f(K)$ est compact. Par conséquent $f(K)$ est un fermé borné de $\mathbb{R}$ par le théorème de Borel-Lebesgue 6.123. Vu que $f(K)$ est borné, la fonction f est bornée.

De plus $f(K)$ étant fermé, son infimum est un minimum et son supremum est un maximum : il existe $x \in K$ tel que $f(x) = \sup f(K)$ et il existe $y \in K$ tel que $f(y) = \inf f(K)$. □

Théorème 6.127 (Théorème de Bolzano-Weierstrass)**.**
Toute suite contenue dans un compact de $\mathbb{R}^m$ admet une sous-suite convergente.

Ce théorème est un cas particulier du théorème 6.147 qui donne pour tout espace métrique, l'équivalence entre la compacité et la compacité séquentielle.

Démonstration. Nous rappelons qu'une partie compacte de $\mathbb{R}^n$ est fermée et bornée par le théorème de Borel-Lebesgue 6.123.

Soit (x_n) une suite contenue dans une partie bornée de $\mathbb{R}^m$. Considérons (a_n), la suite réelle des premières composantes des éléments de (x_n) : pour chaque $n \in \mathbb{N}$, le nombre a_n est la première

composante de x_n. Étant donné que la suite (x_n) est bornée, il existe un M tel que $\|x_n\| < M$. La croissance de la fonction racine carré donne

$$|a_n| \leqslant \|x_n\| \leqslant M. \tag{6.99}$$

La suite (a_n) est donc une suite réelle bornée et donc contient une sous-suite convergente. Soit a_{I_1} une sous-suite convergente de (a_n). Nous considérons maintenant x_{I_1}, c'est à dire la suite de départ dont on a enlevé tous les éléments qu'il faut pour qu'elle converge en ce qui concerne la première composante.

Si nous considérons la suite b_{I_1} des *secondes* composantes de x_{I_1}, nous en extrayons, de la même façon que précédemment, une sous-suite convergente, c'est à dire que nous avons un $I_2 \subset I_1$ tel que b_{I_2} est convergent. Notons que a_{I_2} est une sous-suite de la (sous) suite convergente x_{I_1}, et donc a_{I_2} est encore convergente.

En continuant ainsi, nous construisons une sous-sous-sous-suite x_{I_3} telle que la suite des *troisième* composantes est convergente. Lorsque nous avons effectué cette procédure m fois, la suite x_{I_m} est une suite dont toutes les composantes convergent, et donc est une suite convergente par la proposition 6.96.

Le tableau suivant donne un petit schéma de la façon dont nous procédons. Les $\bullet$ sont les éléments de la suite que nous gardons, et les $\times$ sont ceux que nous «jetons».

$$\begin{array}{lcccccccccc} x_{\mathbb{N}} & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \dots \\ x_{I_1} & \times & \bullet & \bullet & \times & \bullet & \times & \times & \bullet & \bullet & \bullet & \dots \\ x_{I_2} & \times & \bullet & \times & \times & \bullet & \times & \times & \bullet & \bullet & \times & \dots \\ \vdots & & & & & & & & & & & \\ x_{I_m} & \times & \times & \times & \times & \bullet & \times & \times & \times & \bullet & \times & \dots \end{array} \tag{6.100}$$

La première ligne, $x_{\mathbb{N}}$, est la suite de départ. □

Corollaire 6.128.
Si un suite est croissante et bornée alors elle est convergente.

Démonstration. Nous nommons (x_n) la suite et nous prenons un majorant M. Toute la suite est alors contenue dans le compact $[x_0, M]$, ce qui donne une sous-suite $(x_{\alpha(n)})$ convergente par le théorème de Bolzano-Weierstrass 6.127. Si ℓ est la limite de cette sous-suite alors nous avons $\ell \geqslant x_n$ pour tout n.

Pour tout $\epsilon > 0$ il existe K tel que si $n > K$ alors $|\ell - x_{\alpha(n)}| < \epsilon$. Vu que ℓ majore la suite nous avons même

$$x_{\alpha(n)} + \epsilon > \ell. \tag{6.101}$$

Vu que la suite est croissante pour tout $m > \alpha(K)$ nous avons $x_m + \epsilon > l$, ce qui signifie $|x_m - \ell| < \epsilon$. □

Nous aurons une version pour les fonctions croissantes et bornées en la proposition 12.88.

Proposition 6.129.
Soient V et W deux espaces vectoriels normés. Soit K, une partie compacte de V, et $f\colon K \to W$, une fonction continue. Alors l'image $f(K)$ est compacte dans W.

Ce résultat est démontré dans un cadre plus général par le théorème 6.77.

Démonstration. Nous allons prouver que $f(K)$ est fermée et bornée.

$\underline{f(K)}$ **est fermé** Nous allons prouver que si (y_n) est une suite convergente contenue dans $f(K)$, alors la limite est également contenue dans $f(K)$. Dans ce cas, nous aurons que l'adhérence de $f(K)$ est contenue dans $f(K)$ et donc que $f(K)$ est fermé. Pour chaque $n \in \mathbb{N}$, le vecteur y_n appartient à $f(K)$ et donc il existe un $x_n \in K$ tel que $f(x_n) = y_n$. La suite (x_n) ainsi construite est une suite dans le fermé K et possède donc une sous-suite convergente

(proposition 6.127). Notons (x'_n) cette sous-suite convergente, et a sa limite : $\lim(x'_n) = a \in K$. Le fait que la limite soit dans K provient du fait que K est fermé.

Nous pouvons considérer la suite $f(x'_n)$ dans W. Cela est une sous-suite de la suite (y_n), et nous avons $\lim f(x'_n) = a$ parce que f est continue. Par conséquent nous avons

$$f(a) = \lim f(x'_n) = \lim y_n. \tag{6.102}$$

Cela prouve que la limite de (y_n) est dans $f(K)$ et par conséquent que $f(K)$ est fermé.

$f(K)$ est borné Si $f(K)$ n'est pas borné, nous pouvons trouver une suite (x_n) dans K telle que

$$\|f(x_n)\|_W > n \tag{6.103}$$

Mais par ailleurs, l'ensemble K étant compact (et donc fermé), nous avons une sous-suite (x'_n) qui converge dans K. Disons $\lim(x'_n) = a \in K$.

Par la continuité de f nous avons alors $f(a) = \lim f(x'_n)$, et donc

$$|f(a)| = \lim |f(x'_n)|. \tag{6.104}$$

La suite $f(x'_n)$ est alors une suite bornée, ce qui n'est pas possible au vu de la condition (6.103) imposée à la suite de départ (x_n).

□

Corollaire 6.130.
Si $f\colon K \to \mathbb{R}$ est une application continue où K est une partie compacte d'un espace vectoriel normé, alors $f(K)$ est borné.

Démonstration. En effet, la proposition 6.129 montre que $f(K)$ est compact et donc borné. □

6.10.5 Connexité par arcs

Définition 6.131.
*Le sous ensemble $A \subset \mathbb{R}^n$ est **connexe par arcs** si pour tout $x, y \in A$, il existe un chemin*[28] *contenu dans A les reliant, c'est-à-dire une application continue*

$$\gamma : [0,1] \to \mathbb{R}^n \text{ tel que } \gamma(0) = x \text{ et } \gamma(1) = y$$

avec $\gamma(t) \in A$ pour tout $t \in [0,1]$.

La connexité d'un ensemble n'implique pas sa connexité par arc. Il suffit pour cela de prendre un ensemble constitué de deux connexes reliés par un chemin de longueur infinie (le graphe d'une fonction de type $\sin(1/x)$ par exemple).

6.10.6 Topologie de la droite réelle complétée

Nous introduisons l'ensemble $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$. À présent les symboles $+\infty$ et $-\infty$ n'ont aucune signification particulière ; il s'agit seulement de deux éléments que nous ajoutons à $\mathbb{R}$ pour former un ensemble que nous notons $\bar{\mathbb{R}}$.

Pas plus tard qu'immédiatement nous leur donnons une signification en définissant une topologie sur $\bar{\mathbb{R}}$. Les ouverts sur $\bar{\mathbb{R}}$ sont

(1) tous les ouverts de $\mathbb{R}$,

(2) les intervalles de la forme $]-\infty, a[$ pour tous les $a \in \mathbb{R}$,

(3) les intervalles de la forme $]a, +\infty[$ pour tous les $a \in \mathbb{R}$,

(4) la topologie engendrée par toutes ces parties de $\bar{\mathbb{R}}$.

28. Attention : ici quand on dit *chemin*, on demande que l'application soit continue. Dans de nombreux cours de géométrie différentielle, on demande C^∞. Il faut s'adapter au contexte.

Par construction, les boules de $\mathbb{R}$ et les intervalles $]-\infty, a[$ et $]a, +\infty[$ forment une base de topologie pour $\bar{\mathbb{R}}$.

Si f est une fonction $f: \mathbb{R} \to \mathbb{R}$, que signifie $\lim_{x\to\infty} f(x)$? Il s'agit de considérer la fonction élargie

$$\begin{aligned} \tilde{f}: \bar{\mathbb{R}} &\to \bar{\mathbb{R}} \\ x &\mapsto \begin{cases} f(x) & \text{si } x \in \mathbb{R} \\ 0 & \text{si } x = \pm\infty. \end{cases} \end{aligned} \tag{6.105}$$

Ensuite, c'est la définition topologie usuelle de la limite. Notons que les limites en a ne dépendent pas de la valeur effective de f en a, donc le prolongement par 0 est sans conséquences. Nous pouvions tout aussi bien prolonger par 4.

Le même raisonnement tient pour donner un sens à $\lim_{x\to a} f(x) = \pm\infty$.

6.10.7 Limite pointée ou épointée ?

Nous sommes des intégristes sur ce point. Nous ne savons pas ce qu'est la limite pointée et nous ne voulons pas le savoir ; tant et si bien que nous disons «limite» là où d'autres auraient le scrupule de préciser «épointée». Le fait est que nous voulons

(1) la définition de la continuité est que l'image inverse d'un ouvert est un ouvert

(2) une fonction est continue en a si et seulement si sa limite en a vaut $f(a)$.

Si par incroyable vous lisiez ces lignes dans l'idée de préparer un concours d'enseignement en France, soyez toutefois au courant de l'existence d'un truc qui s'appelle «limite pointée».

6.11 Espaces métriques

6.11.1 Espaces métrisables

Définition 6.132.
Un espace topologique est ***métrisable*** *s'il est homéomorphe à un espace métrique.*

Proposition 6.133.
Si X est un espace topologique dont la topologie est donnée par une famille dénombrable de semi-normes, alors il est métrisable.

6.11.2 Fonctions continues

La propriété suivante donne des caractérisations importantes de la continuité dans le cas des espaces métriques.

Proposition 6.134 (Continuité, ouverts et voisinages et limite[72])**.**
Soit $f: E \to F$ une application entre espaces métriques et $a \in E$. Alors nous avons équivalence entre les choses suivantes :

(1) f est continue en a,

(2) Pour tout voisinage ouvert W de $f(a)$, il existe un voisinage ouvert V de a tel que $f(V) \subset W$.

(3) Pour toute boule $W' = B\big(f(a), \epsilon\big)$, il existe une boule $V' = B(a, \delta)$ telle que $f(V) \subset W$.

(4) $\forall \epsilon > 0$, $\exists \delta > 0$ tel que $f\big(B(a, \delta)\big) \subset B\big(f(a), \epsilon\big)$.

(5) $\lim_{x\to a} f(x) = f(a)$ où la limite est donnée par la définition 6.65,

(6) Pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que $\|x - a\| < \delta$ implique $\|f(x) - f(a)\| < \epsilon$.

La proposition 6.194 nous montrera que ces équivalences tiennent encore lorsque l'espace a une topologie de semi-normes.

Démonstration. L'équivalence (1) ⇔ (2) est la définition 6.69. L'équivalence (3) ⇔ (4) est une simple paraphrase.

Montrons (2) ⇒ (3). Si $W' = B\big(f(a), \delta\big)$, nous avons un voisinage V de a tel que $f(V) \subset W$. L'ensemble V contenant une boule autour de chacun de ses points [29], il en contient un autour de $a : V' = B(a, \delta) \subset V$. A fortiori nous avons $f(V') \subset W$.

Montrons (3) ⇒ (2). Si W est un ouvert autour de $f(a)$, il contient une boule autour de $f(a)$: $B\big(f(a), \epsilon\big) \subset W$. Il existe donc une boule $V' = B(a, \delta)$ telle que $f(V') \subset B\big(f(a), \epsilon\big) \subset W$.

L'équivalence (1) ⇔ (5) est la définition 6.69 de la continuité en un point couplée à l'unicité de la limite due à la proposition 6.68 parce qu'un espace métrique est séparé.

Prouvons (5) ⇒ (6). Soit $\epsilon > 0$ et $V = B\big(f(a), \epsilon\big)$. Étant donné que $f(a)$ est une limite de f pour $x \to a$, il existe un voisinage W de a tel que $f(W) \subset V$. Soit $\delta > 0$ tel que $B(a, \delta) \subset W$; alors si $\|x - a\| < \delta$ nous avons $x \in B(x, \delta) \subset W$ et donc $f(x) \in B\big(f(a), \epsilon\big)$, c'est à dire $\|f(a) - f(x)\| < \epsilon$.

Enfin l'implication (2) ⇒ (5) est une réécriture de la définition de la limite en un point. □

Proposition 6.135.
Une isométrie entre deux espaces métriques est continue.

Démonstration. Soient $f : X \to Y$, une application isométrique et $\mathcal{O}$ un ouvert de Y. Soit $a \in f^{-1}(\mathcal{O})$; si $d(a, b) < r$, alors $d\big(f(a), f(b)\big) < r$ et donc $b \in f^{-1}\big(B(f(a), r)\big)$. Donc autour de chaque point de $f^{-1}(\mathcal{O})$ nous pouvons trouver une boule ouverte contenue dans $f^{-1}(\mathcal{O})$, ce qui prouve que $f^{-1}(\mathcal{O})$ est ouvert. □

Théorème 6.136 (Théorème des fermés emboîtés[73])**.**
Soit (E, d) un espace métrique. Il est complet si et seulement si toute suite décroissante de fermés non vides dont le diamètre tend vers zéro a une intersection qui se réduit à un seul point.

Démonstration. **Condition suffisante** Soit $\{F_n\}_{n\in\mathbb{N}}$ une telle suite de fermés emboités. Si nous choisissons des points $x_n \in F_n$, nous obtenons une suite (x_n) de Cauchy et qui est par conséquent convergente vu que l'espace est par hypothèse complet. De plus, pour chaque $N \geqslant n$, la queue de suite $(x_n)_{n\geqslant N}$ est contenue dans F_N et donc converge vers un élément de F_N (parce que ce dernier est fermé). Donc la limite de (x_n) est dans $\bigcap_{n\in\mathbb{N}} F_n$.

De plus cette intersection a diamètre nul parce que le diamètre de $\bigcap_{n\in\mathbb{N}} F_n$ est majoré par tous les diamètres des F_n, lesquels sont arbitrairement petits par hypothèse. Donc l'intersection est réduite a un point.

Condition nécessaire Soit une suite de Cauchy (x_n). Nous considérons les ensembles

$$F_n = \overline{\{x_i \text{ tel que } i \geqslant n\}}. \tag{6.106}$$

Le fait que la suite soit de Cauchy implique que $\operatorname{diam}(F_n) \to 0$. Par hypothèse, nous avons alors

$$\bigcap_{n\in\mathbb{N}} F_n = \{a\}. \tag{6.107}$$

Pour s'assurer que a est bien la limite de (x_n), il suffit de remarquer que

$$d(x_n, a) \leqslant \operatorname{diam} F_n \to 0. \tag{6.108}$$

□

6.11.3 Caractérisations séquentielles

Définition 6.137.
*Si X est un espace topologique, une fonction $f : X \to \mathbb{R}$ est **séquentiellement continue** en un point a si pour toute suite convergente $x_n \to a$ dans X nous avons $f(x_n) \to f(x)$ dans $\mathbb{R}$.*

29. Cela est le théorème-définition 6.19 des ouverts dans un espace métrique, à ne pas confondre avec le théorème 6.3.

Une fonction continue est séquentiellement continue. Dans les espaces métriques la proposition suivante montre que la réciproque est également vraie et la continuité est équivalente à la continuité séquentielle. Cela n'est cependant pas vrai pour n'importe quel espace topologique.

Proposition 6.138 (Caractérisation séquentielle de la continuité[1])**.**
Le lien entre continuité et continuité séquentielle.

(1) *Une application continue est séquentiellement continue (quels que soient les espaces de départ et d'arrivée).*

(2) *Si X et Y sont des espaces métriques, alors une fonction $f\colon X \to Y$ est continue en un point si et seulement si elle est séquentiellement continue.*

Démonstration. **Sens direct** Supposons que f soit continue en a et considérons une suite $a_k \to a$. Soit V un voisinage de $f(a)$ et W un voisinage de a tel que $f(W) \subset V$ (définition 6.65 de la continuité en un point). Par la convergence $a_k \to a$, il existe N tel que pour tout $k > N$, $a_k \in W$, et donc tel que $f(a_k) \in V$, ce qui donne la continuité séquentielle de f.

Sens réciproque, espaces métriques Nous supposons maintenant que X et Y soient métriques. Si f n'est pas continue en a, il existe $\epsilon > 0$ tel que pour tout $\delta > 0$, il existe x tel que $\|x - a\| \leqslant \delta$ et $\|f(x) - f(a)\| > \epsilon$. Nous considérons un tel ϵ et pour chaque $n \geqslant 1 \in \mathbb{N}$ nous considérons un x_n correspondant à $\delta = \frac{1}{n}$. Cela nous donne une suite $x_n \to a$ dans X mais $\|f(x_n) - f(a)\|$ reste plus grand que ϵ. Cela montre que f n'est pas non plus séquentiellement continue.

□

Les espaces métriques ont une propriété importante que la fermeture séquentielle est équivalente à la fermeture.

Proposition 6.139 (Caractérisation séquentielle d'un fermé)**.**
Soit X un espace métrique et $F \subset X$. L'ensemble F est fermé si et seulement si toute suite contenue dans F et convergeant dans X converge vers un élément de F.

Démonstration. Si F est fermé alors $X\backslash F$ est ouvert. Soit une suite $x_n \xrightarrow{X} x$ contenue dans F. Si A est un ouvert autour de x contenu dans $X\backslash F$ (existence par le théorème 6.3), alors la suite ne peut pas entrer dans A et ne peut donc pas converger vers x.

Dans l'autre sens maintenant. Supposons que $X\backslash F$ ne soit pas ouvert. Alors il existe $x \in X\backslash F$ pour lequel tout voisinage intersecte F. En prenant $x_k \in B(x, \frac{1}{k})$, nous construisons une suite contenue dans F qui converge vers x. □

Proposition 6.140.
Soient deux espaces métriques E et Y. Soit $f\colon E \to Y$ une application séquentiellement continue. Alors f est continue.

Démonstration. Soit $\mathcal{O}$ un ouvert de Y ; nous allons voir que le complémentaire de $f^{-1}(\mathcal{O})$ est fermé dans E. Pour cela nous considérons une suite convergente $x_k \xrightarrow{E} x$ avec $x_k \in \complement f^{-1}(\mathcal{O})$ pour tout k. Nous allons montrer que $x \in \complement f^{-1}(\mathcal{O})$ et la caractérisation séquentielle[30] de la fermeture conclura que $\complement f^{-1}(\mathcal{O})$ est fermé.

Pour tout k, nous avons $f(x_k) \in \complement\mathcal{O}$, mais $\mathcal{O}$ est ouvert et $f(x_k) \xrightarrow{Y} f(x)$ parce que f est séquentiellement continue. Par conséquent $f(x) \in \complement\mathcal{O}$ et $x \in \complement f^{-1}(\mathcal{O})$. □

Proposition 6.141.
Une fonction séquentiellement continue sur un espace métrisable et à valeurs dans un espace métrique est continue.

30. Proposition 6.139.

Démonstration. Soit E un espace métrique et un homéomorphisme $\phi\colon X \to (E,d)$. Nous supposons que $f\colon X \to Y$ est séquentiellement continue. Nous considérons l'application $\tilde{f} = f \circ \phi^{-1}$, c'est à dire

$$\begin{aligned}\tilde{f}\colon E &\to Y\\ a &\mapsto f\big(\phi^{-1}(a)\big).\end{aligned} \tag{6.109}$$

L'application ϕ^{-1} est continue et donc séquentiellement continue. De plus $\tilde{f}$ est séquentiellement continue. En effet si $a_k \xrightarrow{E} a$, alors

$$\tilde{f}(a_k) = f\big(\phi^{-1}(a_k)\big), \tag{6.110}$$

mais ϕ^{-1} est séquentiellement continue, donc $\phi^{-1}(a_k) \xrightarrow{X} \phi^{-1}(a)$, ce qui signifie que $\phi^{-1}(a_k)$ est une suite convergente dans X et donc

$$\lim_{k\to\infty} \tilde{f}(a_k) = \lim_{k\to\infty} f\big(\phi^{-1}(a_k)\big) = f\big(\phi^{-1}(a)\big) = \tilde{f}(a). \tag{6.111}$$

L'application $\tilde{f}$ est donc séquentiellement continue. Mais étant donné que $\tilde{f}$ est définie sur un espace métrique (E) et à valeurs dans un métrique, elle est continue par la proposition 6.140. L'application $f = \tilde{f} \circ \phi$ est donc continue en tant que composée d'applications continues. □

Proposition 6.142.
Si X et Y sont deux espaces métriques et $f,g\colon X \to Y$ sont deux fonctions continues égales sur une partie dense de X alors $f = g$.

Démonstration. Les fonctions f et g sont séquentiellement continues (proposition 6.138(2)). Soit A un ensemble dense dans X sur lequel f et g sont égales, et $x \notin A$. Vu que A est dense, il existe une suite a_n dans A telle que $a_n \to x$. La séquentielle continuité de f et g donnent

$$f(a_n) \to f(x) \tag{6.112a}$$
$$g(a_n) \to g(x), \tag{6.112b}$$

mais pour tout n, $f(a_n) = g(a_n)$. Par unicité de la limite[31] dans Y, $f(x) = g(x)$. □

6.11.4 Distance entre un point et un ensemble

Définition 6.143.
*Si A est une partie de l'espace métrique (X,d) et si $x \in X$, nous disons que la **distance** entre A et x est le nombre*

$$d(x,A) = \inf_{a\in A} d(x,a). \tag{6.113}$$

Proposition 6.144.
Soit (X,d) un espace topologique métrique et F un fermé de X. Nous avons $d(x,F) = 0$ si et seulement si $x \in F$.

Démonstration. Si $x \in F$ alors $d(x,F) = 0$ parce que $d(x,x)$ fait partie de l'ensemble sur lequel nous prenons l'infimum.

Si réciproquement $d(x,F) = 0$, cela signifie que pour tout ϵ, il existe $x_\epsilon \in F$ tel que $d(x_\epsilon, x) \leqslant \psi$. En prenant $\epsilon = 1/k$ nous construisons une suite (x_k) d'éléments dans F vérifiant $d(x_k,x) = \frac{1}{k}$. Cela signifie que $\lim_{k\to\infty} x_k = x$ par la proposition 6.52(1).

Par la caractérisation séquentielle des fermés (un fermé contient les limites de toutes ses suites, proposition 6.139), la suite (x_k) étant dans F, la limite est dans F. Donc $x \in F$. □

31. Proposition 6.68.

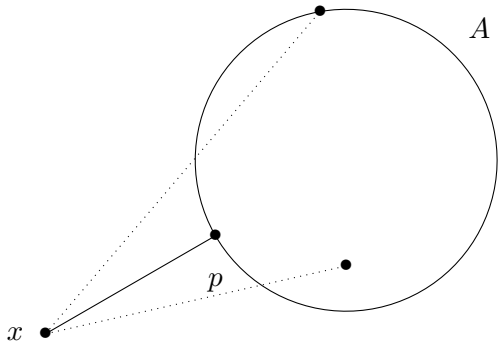


FIGURE 6.4 – La distance entre x et A est donnée par la distance entre x et p. Les distances entre x et les autres points de A sont plus grandes que $d(x,p)$.

6.11.5 Compacité

Lemme 6.145 (de Lebesgue[74])**.**
Soit (X,d) un espace métrique tel que toute suite ait une sous-suite convergente à l'intérieur de l'espace. Si $\{V_i\}$ est un recouvrement par des ouverts de X, alors il existe ϵ tel que pour tout $x \in X$, nous ayons $B(x,\epsilon) \subset V_i$ pour un certain i.

Démonstration. Par l'absurde, nous supposons que pour tout n, il existe un $x_n \in X$ tel que la boule $B(x_n, \frac{1}{n})$ n'est contenue dans aucun des V_i. Ce des x_n nous extrayons une sous-suite convergente (que nous nommons encore (x_n)) et nous posons $x_n \to x$. Pour n assez grand ($\frac{1}{n} < \epsilon$) nous avons $x_n \in B(x,\epsilon)$, donc tous les x_n suivants sont dans le V_i qui contient x. □

Lemme 6.146 ([74])**.**
Soit (X,d) un espace métrique tel que toute suite possède une sous-suite convergente. Pour tout $\epsilon > 0$, il existe un ensemble fini $\{x_i\}_{i\in I}$ tel que les boules $B(x_i,\epsilon)$ recouvrent X.

Démonstration. Soit par l'absurde un $\epsilon > 0$ contredisant le lemme. Il n'existe pas d'ensemble finis autour des points duquel les boules de taille ϵ recouvrent X.

Nous construisons par récurrence une suite ne possédant pas de sous-suites convergente. Le premier terme, x_0 est pris arbitrairement dans X. Ensuite si nous en avons N termes, nous savons que les boules de rayon ϵ et centrées en les points $\{x_i\}_{i=1,\dots,N}$ ne recouvrent pas X. Donc nous prenons x_{N+1} hors de l'union de ces boules.

Ainsi nous avons une suite (x_n) dont tous les termes sont à distance plus grande que ϵ les uns des autres. Une telle suite ne peut pas contenir de sous-suite convergente. Contradiction. □

Théorème 6.147 (Bolzano-Weierstrass[74])**.**
Un espace métrique est compact si et seulement si toute suite admet une sous-suite qui converge à l'intérieur de l'espace.

Une version dédiée à $\mathbb{R}^n$ sera démontrée dans le théorème 6.127.

Démonstration. Soit X un espace métrique compact et (x_n) une suite dans X. Nous considérons la suite de fermés emboités

$$X_n = \overline{\{x_k \text{ tel que } k > n\}}. \tag{6.114}$$

Ce sont des fermés ayant la propriété d'intersection finie non vide, et donc la proposition 6.73 nous dit qu'ils ont une intersection non vide. Un élément de cette intersection est automatiquement un point d'accumulation de la suite.

Nous passons à l'autre sens. Nous supposons que toute suite dans X contient une sous-suite convergente, et nous considérons $\{V_i\}_{i\in I}$, un recouvrement de X par des ouverts. Par le lemme 6.145, nous considérons un ϵ tel que pour tout x, il existe un $i \in I$ avec $B(x,\epsilon) \subset V_i$. Par le lemme 6.146, nous considérons un ensemble fini $\{y_i\}_{i\in A}$ tel que le boules $B(y_i,\epsilon)$ recouvrent X.

Par construction, chacune de ces boules $B(y_i, \epsilon)$ est contenue dans un des ouverts V_i. Nous sélectionnons donc parmi les V_i le nombre fini qu'il faut pour recouvrir les $B(y_i, \epsilon)$ et donc pour recouvrir X. □

Exemple 6.148(Non compacité de la boule unité en dimension infinie)
Le théorème de Bolzano-Weierstrass permet de voir tout de suite que la boule unité n'est pas compacte dans un espace vectoriel de dimension infinie : la suite des vecteurs de base ne possède pas de sous-suites convergentes. △

Le théorème de Bolzano–Weierstrass 6.147 a l'importante conséquence suivante.

Théorème 6.149 (Weierstrass)**.**
Une fonction continue à valeurs réelles définie sur un compact est bornée et atteint ses bornes.

Démonstration. Soit K un compact et $f\colon K \to \mathbb{R}$ une fonction continue. Nous désignons par A l'ensemble des valeurs prises par f sur K :

$$A = f(K) = \{f(x) \text{ tel que } x \in K\}. \tag{6.115}$$

Nous considérons le supremum $M = \sup A = \sup_{x\in K} f(x)$ avec la convention comme quoi si A n'est pas borné supérieurement, nous posons $M = \infty$ (voir définition 6.106).

Nous allons maintenant construire une suite (x_n) de deux façons différentes suivant que $M = \infty$ ou non.

(1) Si $M = \infty$, nous choisissons, pour chaque $n \in \mathbb{N}$, un $x_n \in K$ tel que $f(x_n) > n$. Cela est certainement possible parce que si A n'est pas borné, nous pouvons y trouver des nombres aussi grands que nous voulons.
(2) Si $M < \infty$, nous savons que pour tout ε, il existe un $y \in A$ tel que $y > M - \varepsilon$. Pour chaque n, nous choisissons donc $x_n \in K$ tel que $f(x_n) > M - \frac{1}{n}$.

Quel que soit le cas dans lequel nous sommes, la suite (x_n) est une suite dans K qui est compact, et donc nous pouvons en extraire une sous-suite convergente à l'intérieur de K par le théorème de Bolzano-Weierstrass6.147. Afin d'alléger la notation, nous allons noter (x_n) la sous-suite convergente. Nous avons donc

$$x_n \to x \in K. \tag{6.116}$$

Par la proposition 6.138, nous avons que f prend en x la valeur

$$f(x) = \lim_{n\to\infty} f(x_n). \tag{6.117}$$

Donc $f(x) < \infty$. Évidement, si nous avions été dans le cas où $M = \infty$, la suite x_n aurait été choisie pour avoir $f(x_n) > n$ et donc il n'aurait pas été possible d'avoir $\lim_{n\to\infty} f(x_n) < \infty$. Nous en concluons que $M < \infty$, et donc que f est bornée sur K.

Afin de prouver que f atteint sa borne, c'est à dire que $M \in A$, nous considérons les inégalités

$$M - \frac{1}{n} < f(x_n) \leqslant M. \tag{6.118}$$

En passant à la limite $n \to \infty$, ces inégalités deviennent

$$M \leqslant f(x) \leqslant M, \tag{6.119}$$

et donc $f(x) = M$, ce qui prouve que f atteint sa borne M au point $x \in K$. □

Lemme 6.150.
Soit A_n une suite décroissante de fermés dans un espace métrique[32] *compact K. Alors*

$$C = \bigcap_{n\in\mathbb{N}} A_n \tag{6.120}$$

est non vide.

32. L'hypothèse métrique provient de l'utilisation de Bolzano-Weierstrass, lequel est vrai pour les espaces séquentiellement compacts, dont les espaces métriques.

Démonstration. Soit (x_n) une suite dans K telle que $x_n \in A_n$. La suite étant contenue dans A_1, et A_1 étant compact (lemme 6.75), elle possède une sous suite $(y_n = x_{\sigma_1(n)})$ convergente dont la limite est dans A_1 par le théorème de Bolzano-Weierstrass 6.147. Une queue de la suite y_n est dans A_2 et nous considérons donc une sous suite convergente dans A_2 donnée par

$$z_n = y_{\sigma_2(n)} = x_{\sigma_1\sigma_2(n)}. \tag{6.121}$$

En continuant ainsi nous construisons une suite convergente dans A_k. Nous considérons enfin la suite

$$y_n = x_{\sigma_1\ldots\sigma_n(n)}. \tag{6.122}$$

Pour tout k, une queue de cette suite est une sous suite de $x_{\sigma_1\ldots\sigma_k(n)}$ et par conséquent cette suite converge dans A_k. La limite de cette suite est donc dans l'intersection demandée. □

Remarque 6.151.
Cette propriété est fausse pour les ouverts. Par exemple

$$\bigcap_{n>1}]0, \frac{1}{n}[= \varnothing. \tag{6.123}$$

Lemme 6.152.
Si K est un compact dans un espace métrique et F un fermé disjoint de K, alors $d(K, F) > 0$.

Démonstration. Le fonction

$$\begin{aligned} K &\to \mathbb{R} \\ x &\mapsto d(x, F) \end{aligned} \tag{6.124}$$

est une fonction continue sur K, et donc atteint son minimum par le théorème de Weierstrass 6.149. Soit $x_0 \in K$ un point de K qui réalise ce minimum. Si $d(x_0, F) = 0$, alors on aurait une suite (x_n) dans F qui convergerait vers x_0, mais F étant fermé cela signifierait que x_0 serait dans F, ce qui contredirait l'hypothèse que F et K sont disjoints. □

Proposition 6.153 ([74])**.**
Une isométrie d'un espace métrique compact sur lui-même est une bijection.

Démonstration. Soit X un espace métrique compact et $f\colon X \to X$ une isométrie. Le fait que f soit injective est obligatoire (sinon il y a des images dont la distance est nulle). Il faut montrer que f est surjective.

Soit $x \in X$ hors de $f(X)$. Le lemme 6.152 appliqué au fermé $\{x\}$ et au compact $f(K)$ donne un $r > 0$ tel que

$$d\big(x, f(K)\big) > r. \tag{6.125}$$

Soit la suite $u_n = f^n(x)$; c'est une suite dans K et possède donc une sous-suite convergente (Bolzano-Weierstrass6.147) que l'on nomme (y_n). Vu que f est une isométrie,

$$d(y_n, y_{n+1}) = d(x, y_m) > r \tag{6.126}$$

pour un certain $m \leqslant n + 1$. Ce la signifie que pour tout n, nous avons $d(y_n, y_{n+1}) > r$, ce qui contredit le fait que la suite (y_n) converge. □

Proposition 6.154.
Soit (X, d) un espace métrique compact et (u_n) une suite de X telle que

$$\lim_{n\to\infty} d(u_n, u_{n+1}) = 0. \tag{6.127}$$

Alors l'ensemble des points d'accumulation de (u_n) est connexe.

Démonstration. Nous notons Γ l'ensemble des points d'accumulation de la suite.

Γ est compact Nous notons $A_p = \{u_n \text{ tel que } n \geqslant p\}$ et nous avons

$$\Gamma = \bigcap_{p\in\mathbb{N}} \overline{A_p} \tag{6.128}$$

parce que si $x \in \Gamma$, alors pour tout n, il existe $m > n$ tel que $x_m \in B(x, \epsilon)$, et donc tel que $x \in B(x_m, \epsilon)$. Donc pour tout ϵ et pour tout p, l'intersection $B(x, \epsilon) \cap A_p$ est non vide.

En tant qu'intersection de fermés, Γ est fermé (lemme 6.2). En tant que fermé dans un compact, Γ est compact (lemme 6.75).

Recouvrement par deux compacts Supposons que Γ ne soit [33] pas connexe. Nous pouvons alors considérer S et O, deux ouverts disjoints recouvrant Γ et intersectant tout deux Γ. Nous posons alors

$$A = S \cap \Gamma \tag{6.129a}$$
$$B = O \cap \Gamma, \tag{6.129b}$$

et nous avons évidemment $\Gamma = A \cup B$. Montrons que A est fermé (B le sera aussi par le même raisonnement). Soit une suite d'éléments de $S \cap \Gamma$ convergent dans X. Alors la limite est dans $\bar{\Gamma} = \Gamma$ et donc elle est donc O ou S, mais elle est certainement dans $\bar{S}$. Cependant $\bar{S}$ n'intersecte pas O. En effet si $x \in \bar{S} \cap O$, alors tout voisinage de x intersecterait S, mais il y a des voisinages de x étant inclus dans O parce que O est ouvert ; cela donnerait une intersection entre O et S, ce qui est impossible. Donc la limite n'est pas dans O et donc elle est dans S. Au final la limite est dans $S \cap \Gamma$, ce qui prouve son caractère fermé.

Comme d'habitude, $\Gamma \cap S$ est compact parce que fermé dans un compact.

Décomposition en trois morceaux Vu que A et B sont des compacts disjoints, nous avons $d(A, B) = \alpha > 0$ pour un certain α par le lemme 6.152. Nous notons

$$A' = \{x \in X \text{ tel que } d(x, A) < \frac{\alpha}{3}\} \tag{6.130a}$$
$$B' = \{x \in X \text{ tel que } d(x, B) < \frac{\alpha}{3}\} \tag{6.130b}$$

Nous avons $A' = \bigcup_{x\in A} B(x, \frac{\alpha}{3})$ et donc en tant qu'union d'ouverts, A' est ouvert (définition de la topologie). Même chose pour B'.

Enfin nous notons

$$K = X\backslash(A' \cup B') \tag{6.131}$$

qui est fermé en tant que complémentaire d'ouvert, et donc compact. Étant donné que $A \subset A'$ et $B \subset B'$, nous avons $K \cap \Gamma = \varnothing$.

L'idée est maintenant de montrer que K contient un point d'accumulation de (u_n).

Sous-suites de (u_n) L'hypothèse sur la suite (u_n) nous indique qu'il existe un N_0 tel que $\forall n \geqslant N_0$,

$$d(u_n, u_{n+1}) < \frac{\alpha}{3}. \tag{6.132}$$

Soit $N > N_0$ et $x_0 \in A$. Étant donné que x_0 est point d'accumulation de la suite, il existe $n_1 > N$ tel que $d(x_0, u_{n_1}) < \frac{\alpha}{3}$. Même chose dans B : nous prenons $y_0 \in B$ et un naturel $n_2 > n_1$ tel que $d(y_0, u_{n_2}) < \frac{\alpha}{3}$. Nous avons $u_{n_1} \in A'$ et $u_{n_2} \in B'$.

Soit n_0 le plus petit naturel supérieur à n_1 tel que $u_{n_0} \notin A'$. Cela existe parce que $u_{n_2} \in B'$ et $B' \cap A' = \varnothing$, mais n_0 n'est pas n_2 lui-même parce que $d(A', B') \geqslant \frac{\alpha}{3}$ alors que nous considérons $n_0, n_1, n_2 > N_0$ et donc pour tous les i entre n_1 et n_2 (compris), $d(u_i, u_{i+1}) < \frac{\alpha}{3}$. Notons qu'ici le strict dans la condition (6.132) est important. Nous avons donc $N_0 < n_1 < n_0 < n_2$.

33. est-ce qu'il faut vraiment un subjonctif ici ?

Nous allons maintenant montrer que u_{n_0} est dans K. C'est fait pour : il est loin en même temps de A' et de B'. En utilisant l'inégalité triangulaire à l'envers, nous avons

$$\begin{aligned} d(u_{n_0}, B) &\geqslant d(u_{n_0-1}, B) - d(u_{n_0-1}, u_{n0}) \\ &\geqslant d(A, B) - d(u_{n_0-1}, A) - d(u_{n_0-1}, u_{n_0}) \\ &\geqslant \alpha - \frac{\alpha}{3} - \frac{\alpha}{3} \\ &= \frac{\alpha}{3}. \end{aligned} \tag{6.133}$$

Pour la dernière inégalité nous avons utilisé le fait que u_{n_0-1} n'est pas dans A'. Bref, nous avons montré que u_{n_0} n'est pas dans B' (dans la définition de ce dernier nous avons bien une inégalité stricte). Vu que par définition u_{n_0} n'est pas non plus dans A', nous avons $u_{n_0} \in K$.

Nous avons montré jusqu'à présent que pour tout $N \geqslant N_0$, il existe un $n_0 \geqslant N$ tel que $u_{n_0} \in K$. Cela nous construit donc une sous-suite (v_n) de (u_n) contenue dans K. En tant que suite dans le compact K, la suite (v_n) admet un point d'accumulation dans K. Ce point est également point d'accumulation de la suite (u_n) complète, ce qui donne un point d'accumulation de (u_n) dans K et donc une contradiction.

Nous concluons que Γ est connexe. □

Encore une petite conséquence sans ambition du théorème de Bolzano-Weierstrass.

Proposition 6.155.
Si (x_n) est une suite dans un compact telle que toute sous-suite convergente ait le même point x comme limite. Alors la suite entière converge vers x.

Démonstration. Supposons que ce ne soit pas le cas. Alors il existe un ϵ tel que pour tout $N > 0$, il existe $n > N$ avec $d(x_n, x) > \epsilon$. Cela nous donne une sous-suite de (x_n) composée d'éléments tous à une distance de x supérieure à ϵ. Nous la nommons (y_n) ; c'est une suite dans un compact qui admet donc une sous-suite convergente (et une telle sous-suite est une sous-suite de (x_n)) dont la limite devrait être x, mais c'est impossible par construction. □

Lemme-définition 6.156.
Soi Ω un ouvert dans un espace métrique E. Il existe une suite (K_n) de compacts tels que

(1) $K_n \subset \Omega$
(2) $\bigcup_{n=0}^{\infty} K_n = \Omega$
(3) $K_n \subset \text{Int}(K_{n+1})$.

Une telle suite de compacts vérifie alors

(1) Il existe δ_n tel que pour tout $z \in K_n$, $B(z, \delta_n) \subset K_{n+1}$.
(2) Tout compact de Ω est inclus à $\text{Int}(K_n)$ pour un certain n.

*Une telle suite de compacts est une **suite exhaustive** de compacts pour Ω.*

Démonstration. Nous considérons les ensembles

$$V_n = \{z \in E \text{ tel que } |z|\} \cup \bigcup_{a \notin \Omega} B(a, \frac{1}{n}), \tag{6.134}$$

et nous définissons $K_n = \complement V_n$. Vérifions que ces ensembles vérifient tout ce qu'il faut.

(1) Si $a \notin \Omega$ alors a est dans tous les V_n et donc dans aucun des K_n ; nous avons donc bien $K_n \subset \Omega$.
(2) Si $z \in \Omega$ alors nous prenons $n_1 > |z|$ puis n_2 tel que $B(z, \frac{1}{n_2}) \subset \Omega$. Alors $z \in K_n$ avec $n > \max(n_1, n_2)$.

(3) Une chose à comprendre est que si $z \in K_n$, alors $d(z, \complement\Omega) \geqslant \frac{1}{n}$. Du coup si nous prenons δ tel que
$$\frac{1}{n+1} < \delta < \frac{1}{n} \tag{6.135}$$
alors $B(z, \delta) \subset K_{n+1}$.

(4) Enfin, les K_n sont tous compacts. En effet ils sont bornés parce que $K_n \subset B(0, n)$ et ensuite K_n est fermé en tant que complémentaire d'un ouvert (V_n est ouvert en tant qu'union d'ouverts).

Nous passons maintenant aux propriétés, qui sont indépendantes de la façon dont nous avons construit les K_n vérifiant les conditions.

(1) Nous pouvons considérer la fonction $K_n \to \mathbb{R}$ donnée par $z \mapsto d(z, \complement K_{n+1})$. Vu que $K_n \subset \text{Int}(K_{n+1})$, c'est une fonction (continue sur le compact K_n) prenant des valeurs strictement positives. Elle a donc un minimum strictement positif. Si δ_n est plus petit que ce minimum nous avons $B(z, \delta_n) \subset K_{n+1}$ pour tout $z \in K_n$.

(2) D'abord nous avons $\Omega = \bigcup_{n=0}^{\infty} \text{Int}(K_n)$. En effet nous avons
$$\Omega = \bigcup_{n=0}^{\infty} K_n \subset \bigcup_{n=0}^{\infty} \text{Int}(K_{n+1}) \subset \bigcup_{n=0}^{\infty} \text{Int}(K_n). \tag{6.136}$$
L'inclusion dans l'autre sens est facile.
Soit K compact dans Ω. Vu que Ω est l'union des $\text{Int}(K_n)$, nous avons
$$K \subset \bigcup_{n=0}^{\infty} \text{Int}(K_n). \tag{6.137}$$
Cela donne à K un recouvrement par des ouverts dont nous pouvons extraire un sous-recouvrement fini par compacité. Les K_n étant croissants, du recouvrement fini, il suffit de prendre le plus grand (disons K_m) et nous avons $K \subset \text{Int}(K_m)$.

□

Notons qu'avec la suite de K_n telle que construite, le dernier point est réglé en prenant
$$\frac{1}{n+1} < \delta_n < \frac{1}{n}. \tag{6.138}$$

Théorème 6.157 (Tykhonov)**.**
Un produit quelconque d'espaces métriques non vides est compact si et seulement si chacun de ses facteurs est compact.

Nous n'allons donner la preuve que dans le cas d'un produit fini dans le théorème 6.162.

6.11.6 Ensembles enchaînés

Soit (x, d) un espace métrique.

Définition 6.158.
*Une ϵ-**chaîne** joignant les points a et b de X est une suite finie $(u_0, \dots, u_n)$ dans X telle que $u_0 = a$, $u_n = b$ et pour tout $0 \leqslant i \leqslant n-1$ nous avons $d(u_n, u_{n+1}) \leqslant \epsilon$.*

*Une partie A de X est **bien enchaînée** si pour tout $\epsilon > 0$ et pour tout $a, b \in A$, il existe une ϵ-chaîne joignant a et b dans A.*

Les rationnels dans $\mathbb{R}$ sont bien enchaînés.

Proposition 6.159.
Un espace connexe est bien enchaîné.

Proposition 6.160.
La fermeture d'un ensemble bien enchaîné dans un espace métrique compact (X, d) est connexe.

Démonstration. Soit $A \subset X$ un ensemble bien enchaîné, et soient $a, b \in \bar{A}$. Nous construisons une suite (u_k) dans A de la façon suivante. Pour chaque $n > 0$ nous prenons $a' \in B(a, \frac{1}{n}) \cap A$ et $b' \in B(b, \frac{1}{n}) \cap A$. Ensuite nous considérons une $\frac{1}{n}$-chaîne $\{v_i^{(n)}\}_{i \in I_n}$ dans A entre a' et b'. Ici l'ensemble I_n est fini. La suite (u_k) est simplement construite en mettant bout à bout les éléments $v_i^{(n)}$.

La suite ainsi construite est une suite dans A admettant a et b comme points d'accumulation (les autres points d'accumulation sont également dans $\bar{A}$) et telle que $\lim_{k\to\infty} d(u_k, u_{k+1}) = 0$. Par conséquent la proposition 6.154 nous dit que l'ensemble des points d'accumulation de (u_k) est connexe dans X. Nous le notons $C_{a,b}$.

Si nous fixons $a \in \bar{A}$, alors nous avons

$$\bigcup_{x \in \bar{A}} C_{a,x} = \bar{A}. \tag{6.139}$$

Vu que le membre de gauche est une union de connexes, c'est un connexe par la proposition 6.89. □

En particulier, un espace métrique compact est connexe si et seulement s'il est bien enchaîné.

6.11.7 Produit fini d'espaces métriques

Définition 6.161.
Si $(E_1, d_1), \dots, (E_n, d_n)$ sont des espaces métriques nous mettons la distance suivante sur le produit cartésien $E = E_1 \times \dots \times E_n$:

$$d(x, y) = \max_{i=1,\dots,n} d_i(x_i, y_i). \tag{6.140}$$

Théorème 6.162 ([1])**.**
Un produit fini d'espaces métriques non vides est compact si et seulement si chacun de ses facteurs est compact.

Démonstration. Soient $K_1, \dots, K_n$ des compacts et $K = K_1 \times \dots \times K_n$ le produit muni de sa métrique usuelle de la définition (6.161) (attention : chacun des K_i peut être de dimension infinie) :

$$d(\alpha, \beta) = \max\{d_i(\alpha_i, \beta_i)\} \tag{6.141}$$

où d_i est la distance sur K_i. Si (α_n) est une suite dans K alors la suite $(\alpha_n)_1$ est une suite dans le compact K_1 dont nous pouvons extraire une sous-suite convergente (Bolzano-Weierstrass 6.147). De la sous-suite de α correspondante nous extrayons la sous-suite pour la seconde composante, etc.

En fin de compte nous avons une sous-suite (que nous nommons α également) donc chacune des composantes est convergente. Nous nommons ℓ_k les limites correspondantes. Soit $\epsilon > 0$ pour chaque $k = 1, \dots, n$, il existe $N_k > 0$ tel que si $p > N_k$ alors

$$d\big((\alpha_p)_k - \ell_p\big) \leqslant \epsilon. \tag{6.142}$$

Ici $\alpha_p \in K$ est le p^{e}élément de la suite α et $(\alpha_p)_i \in K_i$ est la i^{e}composante de α_p. En prenant $N = \max_k N_k$ et $n > N$ nous avons

$$d\big(\alpha_n, (\ell_1, \dots, \ell_n)\big) \leqslant \epsilon. \tag{6.143}$$

Par conséquent de la suite (α) nous avons extrait une sous-suite convergente et la partie «réciproque» de Bolzano-Weierstrass nous assure alors que K est compact.

À l'inverse si un des facteurs n'est pas compact (mettons K_1) alors nous prenons un recouvrement $\{\mathcal{O}_i\}_{i\in I}$ de K_1 par des ouverts duquel il est impossible d'extraire un sous-recouvrement fini. Ensuite nous posons

$$\mathcal{P}_i = \mathcal{O}_i \times K_2 \times \ldots \times K_n, \tag{6.144}$$

qui est un recouvrement de K par des ouverts (de K) d'où aucun sous-recouvrement fini ne peut être extrait. □

Pour la culture générale, il y a bien entendu moyen de faire des produits dénombrables et pire d'espaces métriques.

Définition 6.163 ([69])**.**
Soient (E_n, d_n) des espaces métriques. Sur l'ensemble produit $E = \prod_{i=1}^{\infty} E_i$ nous définissons la métrique

$$d(x,y) = \sum_{k=1}^{\infty} \frac{1}{2^k} d'_k(x_i, y_i) \tag{6.145}$$

où $d'_i = \min(d_i, 1)$.

On peut montrer que ce d est bien une distance et que (E, d) devient un espace métrique.

Théorème 6.164 (Tykhonov dénombrable[69])**.**
Un produit dénombrable d'espaces métriques non vides est compact si et seulement si chacun de ses facteurs est compact.

Note : ce résultat est encore valable pour un produit quelconque, c'est le théorème de Tykhonov 6.157.

6.12 Ensembles nulle part denses

Nous allons nous limiter au cas de $\mathbb{R}$, mais je crois que ça se généralise sans trop de peine aux espaces en tout cas métriques. Voir aussi la section 6.16 sur les espaces de Baire.

Définition 6.165.
*Un ensemble est dit **nulle part dense** s'il n'est dense dans aucun intervalle.*

*Un ensemble dans $\mathbb{R}$ est de **première catégorie** ou **maigre** s'il est une union dénombrable d'ensembles nulle part dense (c'est à dire d'ensembles denses sur aucun intervalle).*

Théorème 6.166 (Baire[75])**.**
Une réunion dénombrable d'ensembles nulle part denses est d'intérieur vide.

Démonstration. Soit $a \in S$ et $\epsilon > 0$. Nous allons trouver un élément dans $B(a,\epsilon)$ qui n'est pas dans S. Nous commençons par choisir $x_1 \in B(a,\epsilon)$ et $r_1 < \frac{\epsilon}{2}$ tel que

$$B(x_1, r_1) \cap A_1 = \varnothing. \tag{6.146}$$

Ensuite nous choisissons $x_2 \in B(x_1, r_1)$ et $r_2 < \epsilon/4$ tel que $B(x_2, r_2) \subset B(x_1, r_1)$ et $B(x_2, r_2) \cap A_2 = \varnothing$. Notons que $B(x_2, r_2) \cap A_1 = \varnothing$ aussi, par construction.

Par récurrence nous construisons une suite d'éléments x_n et de rayons $r_n < \epsilon/2^n$ tels que

(1) $B(x_n, r_n) \cap A_j = \varnothing$ pour tout $j \leqslant n$,
(2) $\overline{B(x_n, r_n)} \subset B(x_{n-1}, r_{r-1})$.

Cette suite étant de Cauchy (parce que contenue dans des intervalles emboités de rayon décroissant vers zéro), elle converge[34] donc vers un point qui en particulier appartient à $B(a,\epsilon)$. Mais la limite n'est dans aucun des A_n et donc pas dans S. □

34. Par le théorème 6.57

6.13 Encore de la topologie réelle

Dans cette section, nous travaillons dans l'espace $\mathbb{R}^n$ pour un certain naturel n. Nous y définissons la notion d'ouvert et de fermé, qui sont la base de la topologie générale. Notons que ces définitions n'ont de sens que relativement à l'espace ambiant, aussi un ouvert de $\mathbb{R}$ ne sera en général pas un ouvert de $\mathbb{R}^2$: d'une part, il n'y a pas d'inclusion canonique de $\mathbb{R}$ dans $\mathbb{R}^2$ (les ouverts du second ne sont même pas des sous-ensembles du premier) et, d'autre part, les définitions se basent sur la notion de boule de $\mathbb{R}^n$ qui dépend évidemment de la valeur de n (une boule dans $\mathbb{R}$ est un intervalle, dans $\mathbb{R}^2$ c'est un disque, etc.)

6.13.1 Ouverts et fermés

Définition 6.167.
*La **boule ouverte** de centre $x_0 \in \mathbb{R}^n$ et de rayon $r \in \mathbb{R}^+$ est définie par*

$$B(x_0, r) = \{x \in \mathbb{R}^n \text{ tel que } \|x - x_0\| < r\}, \tag{6.147}$$

*tandis que la **boule fermée** de centre x_0 et de rayon r est*

$$\bar{B}(x_0, r) = \{x \in \mathbb{R}^n \text{ tel que } \|x - x_0\| \leqslant r\}; \tag{6.148}$$

la différence est que l'inégalité dans la première est stricte.

Définition 6.168.
*Une partie A de $\mathbb{R}^n$ est **ouverte** si pour tout $a \in A$ il existe $r > 0$ tel que $B(a, r) \subset A$. Une partie est donc ouverte lorsqu'elle contient une boule autour de chacun de ses éléments.*

Cette définition est évidemment à mettre en rapport avec le théorème 6.3.
Le lemme suivant justifie le vocabulaire des définitions 6.167.

Lemme 6.169.
Pour tout $x \in \mathbb{R}^n$ et tout $r > 0$ la boule $B(x, r)$ est ouverte.

Démonstration. Afin de prouver que la boule est ouverte, nous prenons un point $p \in B(x, r)$, et nous allons montrer qu'il existe une boule autour de p qui est contenue dans $B(x, r)$.

Étant donné que $p \in B(x, r)$, nous avons $d(p, x) < r$. Prouvons que la boule $B\big(p, r - d(p, x)\big)$ est contenue dans $B(x, r)$. Pour cela, nous prenons $p' \in B\big(p, r - d(p, x)\big)$, et nous essayons de prouver que $p' \in B(x, r)$. En effet, en utilisant l'inégalité triangulaire,

$$d(x, p') \leqslant d(x, p) + d(p, p') \leqslant d(x, p) + r - d(p, x) = r. \tag{6.149}$$

□

6.13.2 Intérieur, adhérence et frontière

Définition 6.170.
*Soit $A \subset \mathbb{R}^n$ et $x \in \mathbb{R}^n$. Le point x est **intérieur** à A s'il existe une boule autour de x complètement contenue dans A. L'ensemble des points intérieurs à A est noté $\operatorname{Int} A$ ou $\mathring{A}$, de sorte qu'on a précisément*

$$x \in \operatorname{Int} A \overset{def}{\iff} \exists \epsilon > 0 \text{ tel que } B(x, \epsilon) \subset A.$$

Définition 6.171.
*Le point x est dans l'**adhérence** de A si toute boule autour de x intersecte A. L'ensemble de ces points est noté $\operatorname{Adh} A$ ou $\bar{A}$, et on a donc de manière plus précise*

$$x \in \operatorname{Adh} A \overset{def}{\iff} \forall \epsilon > 0, B(x, \epsilon) \cap A \neq \varnothing \tag{6.150}$$

Proposition 6.172.
Pour $A \subset \mathbb{R}^n$, nous avons

$$\text{Int}\, A \subseteq A \subseteq \text{Adh}\, A$$

Définition 6.173.
*La **frontière** ou le **bord** de A est défini par $\partial A = \text{Adh}\, A \backslash \text{Int}\, A$. L'ensemble A est un **ouvert** si $A = \text{Int}\, A$, et c'est un **fermé** si $A = \text{Adh}\, A$.*

Lemme 6.174 (Caractérisation équivalente de la frontière)**.**
Soit un espace topologique X et une partie $S \subset X$. Un point $x \in X$ est dans ∂S si et seulement si tout voisinage de x contient un point de S et un point de S^c.

Démonstration. Supposons que tout voisinage de x contienne un point de S et un point de S^c. Alors $x \in Adh(S)$ (définition 6.171), mais pas dans l'intérieur de S parce que x ne possède pas de voisinage contenu dans S. Donc $x \in \partial S$.

À l'inverse, si $x \in \partial S$ alors x est dans l'adhérence de S et tout voisinage de x contient un point de S. Mais x n'est pas dans l'intérieur de S et tout voisinage de x contient un point qui n'est pas dans S, aka un point de S^c. □

Corollaire 6.175.
Un ensemble et son complémentaire ont même frontière.

Démonstration. Conséquence du lemme 6.174. Les points de $\partial(S^c)$ sont caractérisés par le fait que tout voisinage contient un point de S^c et un point de $(S^c)^c = S$. □

Exemple 6.176
Soit $X = [0, 1]$ muni de la topologie de la distance $|x - y|$ (définition 6.19). Les points 0 et 1 *ne sont pas* dans la frontière de X. En effet une boule ouverte autour de 1 est un ensemble de la forme

$$B(1, r) = \{x \in X \text{ tel que } |x - 1| < r\} =]1 - r, 1] \tag{6.151}$$

où nous avons supposé $r < 1$.

Les points 0 et 1 sont par contre sur la frontière de $[0, 1]$ lorsque cet ensemble est vu comme partie de l'espace métrique $\mathbb{R}$. △

Lemme 6.177 (Passage de douane[76, 77])**.**
Dans un espace topologique, toute partie connexe qui rencontre à la fois une partie A et son complémentaire rencontre nécessairement la frontière de A.

Démonstration. Nommons γ la partie connexe qui intersecte A et A^c. Les ouverts $\text{Int}(A)$ et $X\backslash\bar{A}$ ne peuvent pas recouvrir γ parce que ce sont deux ouverts disjoints alors que γ est connexe (voir la définition 6.84 de la connexité). Donc γ doit contenir des points qui sont dans $\bar{A}$ mais pas dans $\text{Int}(A)$. C'est à dire des points de ∂A. □

On vérifiera que les notations et les dénominations sont cohérentes en prouvant la proposition suivante.

Proposition 6.178.
Pour $\epsilon > 0$,

(1) l'adhérence de $B(x, \epsilon)$ est $\bar{B}(x, \epsilon)$,

(2) l'intérieur de $\bar{B}(x, \epsilon)$ est $B(x, \epsilon)$,

(3) la boule ouverte $B(x, \epsilon)$ est un ouvert,

(4) la boule fermée $\bar{B}(x, \epsilon)$ est un fermé.

Nous avons également les liens suivants entre intérieur, adhérence, ouvert, fermé et passage au complémentaire (noté ^{c}) :

Proposition 6.179.
Si $A \subset \mathbb{R}^n$ et $A^c = \mathbb{R}^n \backslash A$, nous avons

(1) $(\operatorname{Int} A)^c = \operatorname{Adh}(A^c)$ *et* $(\operatorname{Adh} A)^c = \operatorname{Int}(A^c)$,

(2) A est ouvert si et seulement si A^c est fermé,

(3) $\operatorname{Int} A$ *est le plus grand ouvert contenu dans A,*

(4) $\operatorname{Adh} A$ *est le plus petit fermé contenant A,*

Exemple 6.180
Il n'est en général pas vrai que $\overline{A \cap B} = \bar{A} \cap \bar{B}$. Par exemple si $A = [0, 1[$ et $B =]1, 2]$. Dans ce cas, $A \cap B = \varnothing$ alors que $\bar{A} \cap \bar{B} = \{1\}$. △

6.13.3 Point d'accumulation, point isolé

Soit $D \subset \mathbb{R}$. Un point $a \in D$ est **isolé** dans D (relativement à $\mathbb{R}$) s'il existe $\varepsilon > 0$ tel que

$$[a - \varepsilon, a + \varepsilon] \cap D = \{a\}. \tag{6.152}$$

Autrement dit, il existe un intervalle autour de a dans lequel a est le seul élément de D.

Un point $a \in \mathbb{R}$ est un **point d'accumulation** de D si pour tout $\varepsilon > 0$,

$$\Big([a - \varepsilon, a + \varepsilon]\backslash\{a\}\Big) \cap D \neq \varnothing. \tag{6.153}$$

Autrement dit, quel que soit l'intervalle autour de a que l'on considère, le point a n'est pas tout seul dans D.

Exemple 6.181
Prenons $D = [0, 1[\cup]2, 3]$. Cet ensemble n'a pas de points isolés, et l'ensemble de ses points d'accumulation est $[0, 1] \cup [2, 3]$.

Notez que les points 1 et 2 sont des points d'accumulation de D qui ne font pas partie de D. Il est possible d'être un point d'accumulation de D sans être dans D, mais pour être un point isolé dans D, il faut être dans D. △

Exemple 6.182
Soit $D = \{\frac{1}{n}\}_{n\in\mathbb{N}}$. Tous les points de cet ensemble sont des points isolés (vérifier !). Aucun point de D n'est point d'accumulation. Cependant 0 est un point d'accumulation. △

[78]

6.13.4 Limite de suite

Définition 6.183 (Limite d'une suite dans $\mathbb{R}^m$)**.**
Une suite de points (x_n) dans $\mathbb{R}^m$ est dite ***convergente*** *s'il existe un élément $\ell \in \mathbb{R}^m$ tel que*

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ tel que } \forall n \geqslant N,\ \|x_n - \ell\| < \varepsilon. \tag{6.154}$$

Dans ce cas, nous disons que ℓ est la ***limite*** *de la suite (x_n) et nous écrivons* $\lim x_n = \ell$ *ou plus simplement $x_n \to \ell$.*

Notez aussi la similarité avec la définition 6.90.

Remarque 6.184.
Nous n'écrivons pas «$\lim_{n\to\infty} x_n$» parce que, lorsqu'on parle de suites, la limite est *toujours* lorsque n tend vers l'infini. Il n'y a aucun intérêt à chercher par exemple $\lim_{n\to 4} x_n$ parce que cela vaudrait x_4 et rien d'autre.

Ceci est une différence importante avec les limites de fonctions.

Lemme 6.185 (Unicité de la limite)**.**
Il ne peut pas y avoir deux nombres différents qui satisfont à la condition (6.154)*. En d'autres termes, si ℓ et ℓ' sont deux limites de la suite (x_n), alors $\ell = \ell'$.*

Démonstration. Soit $\varepsilon > 0$. Nous considérons N tel que

$$\|x_n - \ell\| < \varepsilon \tag{6.155}$$

pour tout $n \geqslant N$, et $N' > 0$ tel que

$$\|x_n - \ell'\| < \epsilon \tag{6.156}$$

pour tout $n > N'$. Maintenant, nous prenons n plus grand que N et N' de telle façon que les deux équations pour x_n soient vérifiées en même temps. Alors

$$\|\ell - \ell'\| = \|\ell - x_n + x_n - \ell'\| \leqslant \|\ell - x_n\| + \|x_n - \ell'\| < 2\varepsilon. \tag{6.157}$$

Cela prouve que $\|\ell - \ell'\| = 0$. □

Le théorème de Bolzano-Weierstrass 6.147 dit que dans le cas métrique, la compacité séquentielle est équivalente à la compacité.

6.14 Application réciproque

Définition 6.186 (injection, surjection, bijection)**.**
Soient des ensembles A et B ainsi qu'une application $f\colon A \to B$.

*(1) La fonction f est **injective** si $f(x_1) = f(x_2)$, implique $x_1 = x_2$.*

*(2) La fonction f est **surjective** si tous les éléments de B sont atteins, c'est à dire si pour tout $y \in B$ il existe $x \in A$ tel que $f(x) = y$.*

*(3) La fonction f est une **bijection** entre A et B si elle est injective et surjective, c'est à dire si pour tout $y \in B$ il existe un unique $x \in A$ tel que $f(x) = y$.*

La surjection et l'injection sont des propriétés bien différentes qu'il convient de prouver séparément. De plus une même «formule» peut définir une application injective, surjective, bijective ou non selon le domains sur laquelle nous la considérons.

Définition 6.187.
*Soit une bijection $f\colon A \to B$. L'**application réciproque** de f est la fonction*

$$\begin{aligned} f^{-1}\colon B &\to A \\ y &\mapsto \text{le } x \in A \text{ tel que } f(x) = y. \end{aligned} \tag{6.158}$$

Plus généralement si $f\colon X \to Y$ est une application quelconque et si $S \subset Y$ nous notons

$$f^{-1}(S) = \{x \in X \text{ tel que } f(x) \in S\}, \tag{6.159}$$

et dans le cas où S est réduit à un unique élément y, nous notons $f^{-1}(y)$ au lieu de $f^{-1}\big(\{y\}\big)$. Si de plus $f^{-1}(S)$ est un singleton x, nous noterons $f^{-1}(S) = x$ et non $f^{-1}(S) = \{x\}$.

Les plus acharnés parmi les lecteurs se rendront compte de la différence ontologique fondamentale entre x et $\{x\}$.

Proposition 6.188.
Soit $f\colon A \subset \mathbb{R}^n \to B \subset \mathbb{R}^m$ une bijection continue. Si A est compact, alors $f^{-1}\colon B \to A$ est continue.

Proposition 6.189.
Soient I un intervalle dans $\mathbb{R}$ et $f\colon I \to \mathbb{R}$ une fonction continue strictement monotone. Alors la fonction réciproque $f^{-1}\colon f(I) \to \mathbb{R}$ est continue sur l'intervalle $f(I)$.

6.15 Topologie des semi-normes

Les principaux espaces topologiques construit avec des semi-normes seront les espaces de fonctions de la définition 28.3. Nous verrons également la topologie $*$-faible sur $\mathscr{D}'(\Omega)$ en la définition 28.10.

Définition 6.190.
*Si E est un espace vectoriel, une **semi-norme** sur E est une application $p\colon E \to \mathbb{R}$ telle que*

(1) $p(x) \geqslant 0$,

(2) $p(\lambda x) = |\lambda| p(x)$

(3) $p(x + y) \leqslant p(x) + p(y)$.

Lemme 6.191 ([79])**.**
Si p est une semi-norme nous avons

$$|p(x) - p(y)| \leqslant p(x - y). \tag{6.160}$$

Démonstration. Nous avons d'une part $p(x + h) \leqslant p(x) + p(h)$ et d'autre part $p(x) \leqslant p(x + h) + p(-h) = p(x + h) + p(h)$. En isolant $p(x + h) - p(x)$ dans chacune ce des deux inégalités,

$$-p(h) \leqslant p(x + h) - p(x) \leqslant p(h) \tag{6.161}$$

ou encore

$$|p(x + h) - p(x)| \leqslant p(h) \tag{6.162}$$

qui donne le résultat demandé en posant $h = y - x$. □

Soit une famille $(p_i)_{i\in I}$ de semi-normes sur E. Nous construisons alors une topologie sur E de la façon suivante.

Définition 6.192 (Topologie et semi-normes[80, 81])**.**
*Pour tout J fini dans I nous définissons les **boules ouvertes***

$$B_J(x, r) = \{y \in E \text{ tel que } p_j(y - x) < r \, \forall j \in J\}. \tag{6.163}$$

*La **topologie** sur E donnée par la famille de semi-norme est définie en disant que $\mathcal{O} \subset E$ est ouvert si et seulement si chaque point de $\mathcal{O}$ est dans une boule contenue dans $\mathcal{O}$.*

Proposition 6.193.
Une suite (x_n) dans E converge vers x au sens de la topologie des semi-normes si et seulement si pour tout $i \in I$,

$$p_i(x - x_n) \to 0. \tag{6.164}$$

Démonstration. Si la suite (x_n) converge[35] vers x, alors pour tout ouvert $\mathcal{O}$ autour de x, il existe un N tel que si $n \geqslant N$, alors $x_n \in \mathcal{O}$. En particulier pour tout j et pour tout $\epsilon > 0$, il doit exister un $n \geqslant N_j$ tel que $x_n \in B_j(x, \epsilon)$.

Voyons l'implication inverse. Soit $\epsilon > 0$. Pour tout $i \in I$, il existe un N_i tel que $n \geqslant N_i$ implique $p_i(x - x_n) \leqslant \epsilon$. Si $\mathcal{O}$ est un ouvert, il doit contenir une boule du type $B_J(x, r)$ pour un certain ensemble fini $J \subset I$.

En prenant $N = \max\{N_j \text{ tel que } j \in J\}$, nous avons $p_j(x - x_n) \leqslant \epsilon$ pour tout j et donc $x_n \in B_J(x, r)$. □

La proposition suivante est un vulgaire plagiat de la proposition 6.134.

Proposition 6.194.
Soit une application $f\colon \mathbb{R} \to (E, p_i)_{i\in I}$. Nous avons équivalence entre

35. Définition 6.4.

(1) la fonction f est continue en $t_0 \in \mathbb{R}$,

(2) si W est un voisinage ouvert de $f(t_0)$ il existe un voisinage ouvert V de t_0 (dans $\mathbb{R}$) tel que $f(V) \subset W$,

(3) pour tout $i \in I$ et $\epsilon > 0$ il existe $\delta > 0$ tel que

$$f\big(B(t_0, \delta)\big) \subset B_i\big(f(t_0), \epsilon\big). \tag{6.165}$$

Démonstration. L'équivalence (1) $\Leftrightarrow$ (2) est la définition 6.69.

Prouvons (2) $\Rightarrow$ (3). Soit $i \in I$ et $\epsilon > 0$ et considérons la boule $B_i\big(f(t_0), \epsilon\big)$, qui est un ouvert de E contenant $f(t_0)$. Il existe donc un ouvert V autour de t_0 tel que $f(V) \subset B_i\big(f(t_0), \epsilon\big)$. En particulier V contient une boule $B(t_0, \delta)$ et nous avons

$$f\big(B(t_0, \delta)\big) \subset f(V) \subset B_i\big(f(t_0), \epsilon\big). \tag{6.166}$$

Prouvons (3) $\Rightarrow$ (2). Soit W un ouvert autour de $f(t_0)$. Il existe un $i \in I$ et $\epsilon > 0$ tel que $B_i\big(f(t_0), \epsilon\big) \subset W$. Nous avons alors un $\delta > 0$ tel que

$$f\big(B(t_0, \delta)\big) \subset B_i\big(f(t_0), \epsilon\big) \subset W. \tag{6.167}$$

□

Lorsqu'on a un espace E muni d'une quantité dénombrable de semi-normes $\{p_k\}_{k\in I}$ nous définissons l'écart [36]

$$d(x, y) = \sup_{k\geqslant 1} \min\big\{\frac{1}{k}, p_k(x - y)\big\}. \tag{6.168}$$

Notons que cette écart est invariant par translation au sens où pour tout x, y, h dans E nous avons

$$d(x + h, y + h) = \sup_{k\geqslant 1} \min\big\{\frac{1}{k}, p_k(x - y)\big\} = d(x, y). \tag{6.169}$$

Proposition 6.195 ([79])**.**
La topologie donnée par les boules

$$B(a, r) = \{x \in E \text{ tel que } \forall\, k \leqslant \frac{1}{k}, p_k(x - a) < r\} \tag{6.170}$$

est la même que celle «usuelle» donnée par les semi-normes. En disant «la même» nous entendons le fait que les ouverts sont les mêmes : A est ouvert pour une des deux topologies si et seulement s'il est ouvert pour l'autre.

Démonstration. Pour cette démonstration nous allons préfixer par d les notions topologiques issues des boules (6.170) et par P celle des semi-normes : P-continue, d-ouvert, etc.

D'abord nous avons

$$B(a, r) = \bigcap_{k\leqslant \frac{1}{k}} B_k(a, r). \tag{6.171}$$

Si $\mathcal{O}$ est un d-ouvert, il contient une d-boule autour de chacun de ses points. Or d'après la formule (6.171), une d-boule est une intersection *finie* de P-ouverts et donc est un P-ouvert par définition. Donc $\mathcal{O}$ contient un P-ouvert autour de tous ses points et est donc P-ouvert.

Inversement nous supposons que $\mathcal{O}$ est un P-ouvert. Commençons par prouver que les semi-normes p_k sont d-continues. En effet soient $k \in \mathbb{N}$, $\epsilon \leqslant \frac{1}{k}$ et $x, y \in E$ tels que $d(x, y) \leqslant \epsilon$; nous avons

$$\begin{aligned}
|p_k(y) - p_k(x)| &\leqslant p_k(x - y) && (6.172a)\\
&= \min\{\frac{1}{k}, p_k(x - y)\} && (6.172b)\\
&\leqslant d(x, y) && (6.172c)\\
&\leqslant \epsilon. && (6.172d)
\end{aligned}$$

36. Dans le cas de $E = \mathscr{D}(K)$, la première semi-norme est numérotée à zéro, donc il faudra poser $d(\varphi_1, \varphi_2)$ avec p_{k-1} au lieu de p_k.

Montrons à présent que $\mathcal{O}$ est d-ouverte. Si $a \in \mathcal{O}$, il existe k et r tels que $B_k(a,r) \subset \mathcal{O}$. Soit $x \in B_k(a,r)$ et montrons que si ϵ est suffisamment petit, la d-boule $B(x,\epsilon)$ est inclue à $B_k(a,r)$. Pour cela prenons $y \in B(x,\epsilon)$; nous avons

$$\big|p_k(a-x) - p_k(a-y)\big| \leqslant d(x,y) \leqslant \epsilon. \tag{6.173}$$

Par conséquent le nombre $p_k(a-y)$ est dans l'intervalle

$$p_k(a-x) \pm \epsilon \tag{6.174}$$

et il suffit de prendre $\epsilon < \frac{r-p_k(a-x)}{2}$. □

6.15.1 Espace dual

Définition 6.196.
Soit F, un espace métrique et E un espace topologique vectoriel. Une topologie possible[37] *sur l'espace des applications linéaires $\mathcal{L}(E,F)$ est la* ***topologie $*$-faible*** *qui est la topologie des semi-normes*

$$p_v(T) = \|T(v)\|_F. \tag{6.175}$$

C'est une famille de semi-normes indicées par les éléments de E. Si E est un espace métrique, c'est cette topologie qui sera considérée sur son dual topologique E' des applications continues $E \to \mathbb{R}$.

La proposition suivante indique qu'elle est un peu la topologie de la convergence ponctuelle.

Proposition 6.197.
Soit E un espace muni de la topologie des semi-normes $\{p_i\}_{i\in I}$ et F un espace métrique. Soit une suite (T_n) dans $\mathcal{L}(E,F)$ et $T \in \mathcal{L}(E,F)$. Nous avons $T_n \xrightarrow{*} T$ si et seulement si $T_n(v) \xrightarrow{F} T(v)$ pour tout $v \in E$.

Démonstration. Nous avons équivalence entre les lignes suivantes :

$$T_n \xrightarrow{*} T \tag{6.176a}$$
$$p_v(T_n - T) \to 0 \; \forall v \in E \qquad \text{proposition 6.193} \tag{6.176b}$$
$$\|T_n(v) - T(v)\|_E \to 0 \; \forall v \in E \tag{6.176c}$$
$$T_n(v) \xrightarrow{E} T(v). \tag{6.176d}$$

□

6.15.2 Espace $C^k(\mathbb{R}, E')$

Nous revenons à nos histoires de limites de la définition 6.4.

Proposition 6.198 (Unicité de la limite dans un dual topologique)**.**
Soit E un espace métrique et E' son dual topologique muni de sa topologie de la définition 6.196. Il y a unicité de l'élément de E' vers lequel une fonction $u\colon \mathbb{R} \to E'$ peut converger.

Démonstration. Soit T un élément vers lequel u_t converge lorsque $t \to t_0$. Soit $\epsilon > 0$ et $x \in E$. La boule $B_x(T,\epsilon)$ de E' subordonnée à la norme p_x et centrée en T est un ouvert de E'. Étant donné que u converge vers T il existe $\delta > 0$ tel que $u_t \in B_x(T,\epsilon)$ dès que $|t - t_0| \leqslant \delta$. Nous avons donc, pour tout $x \in E$, la limite (dans $\mathbb{R}$) :

$$\lim_{t\to t_0} u_t(x) = T(x). \tag{6.177}$$

37. C'est, dans l'idée, celle qui sera choisie pour les espaces de distributions, voir la définition 28.10.

Cela prouve que la convergence de u vers T implique l'existence pour tout x de la limite de $u_t(x)$ dans $\mathbb{R}$. Si T' est un autre élément vers lequel u_t converge, nous avons par le même raisonnement que

$$\lim_{t\to t_0} u_t(x) = T'(x). \tag{6.178}$$

Par unicité de la limite dans $\mathbb{R}$ nous devons alors avoir $T(x) = T'(x)$ pour tout x, c'est à dire $T = T'$. □

Proposition 6.199.
Soit une fonction continue $u\colon \mathbb{R} \to E'$. Alors

(1) pour tout $x \in E$ la fonction $t \mapsto u_t(x)$ est continue,

(2) pour tout $x \in E$ nous avons la limite dans $\mathbb{R}$

$$\lim_{y\to t_0} u_t(x) = u_{t_0}(x), \tag{6.179}$$

(3) nous avons la limite dans E'

$$\lim_{t\to t_0} u_t = u_{t_0}. \tag{6.180}$$

Démonstration. Soient $x \in E$ et $\epsilon > 0$. Par la proposition 6.194 la continuité de u donne un $\delta > 0$ tel que

$$u_{B(t_0,\delta)} \subset B_x(u_{t_0}, \epsilon). \tag{6.181}$$

C'est à dire que si $|t - t_0| \leqslant \delta$ nous avons

$$\big|u_{t_0}(x) - u_t(x)\big| < \epsilon, \tag{6.182}$$

ce qui signifie bien que la fonction $t \mapsto u_t(x)$ est continue en tant que fonction $\mathbb{R} \to \mathbb{R}$. Cela est le point (1). Le théorème de limite et continuité dans $\mathbb{R}$ nous donne immédiatement la limite (6.179).

Nous passons à la preuve du point (3). Soit $\mathcal{O}$ un ouvert de E' contenant u_{t_0}. Il existe donc un $i \in I$ et $\epsilon > 0$ tel que $B_i(u_{t_0}, \epsilon) \subset \mathcal{O}$. Étant donné que u est continue, il existe $\delta > 0$ tel que

$$u_{B(t_0,\delta)} \subset B_i(u_{t_0}, \epsilon) \subset \mathcal{O}. \tag{6.183}$$

Cela signifie bien que

$$|t - t_0| \leqslant \delta \Rightarrow u_t \in \mathcal{O}, \tag{6.184}$$

c'est à dire que nous avons la limite $\lim_{t\to t_0} u_t = u_{t_0}$ dans E'. Pour dire cela nous avons utilisé la définition 6.65 de la limite et le résultat d'unicité 6.198. □

Définition 6.200.
*Si nous avons une application $u\colon \mathbb{R} \to E'$ nous considérons sa **dérivée** donnée par la limite*

$$u'_{t_0} = \lim_{t\to t_0} \frac{u_t - u_{t_0}}{t - t_0}. \tag{6.185}$$

Cela est un nouvel élément de E' (pour peu que la limite existe). La fonction $u'\colon \mathbb{R} \to E'$ ainsi définie peut être continue ou non. Cela nous permet de définir les espaces $C^k(\mathbb{R}, E')$ et $C^\infty(\mathbb{R}, E')$.

Une des principales utilisation que nous ferons de ces espaces seront les espaces de fonctions à valeurs dans le distributions tempérées dont nous parlerons dans la section 28.4.

6.16 Espaces de Baire

Définition 6.201.
*Un **espace de Baire** est un espace topologique dans lequel toute intersection dénombrable d'ouverts denses est dense.*

Théorème 6.202 (Théorème de Baire[82])**.**
Les espaces suivants sont de Baire :

(1) les espaces topologiques localement compacts,

(2) les espaces métriques complets (donc ceux de Banach en particulier),

(3) tout ouvert d'un espace de Baire.

Démonstration. **Espaces topologiques localement compact**

Espaces métriques complets Soit (E, d) un espace métrique complet. Soient V un ouvert quelconque de E et U_n une suite d'ouverts denses. Le but est de prouver que l'ensemble $\bigcap_{n\in\mathbb{N}} U_n$ intersecte V. Vu que V est ouvert dans un espace métrique, il contient une boule ouverte et donc une boule fermée B_0 de rayon strictement positif. L'ensemble U_1 est dense et intersecte donc un ouvert contenu dans B_0. L'intersection est un ouvert qui contient alors une boule fermée B_1 de rayon strictement positif. Continuant ainsi nous construisons une suite de fermés emboités B_n telle que

$$\bigcap_{n\in\mathbb{N}} U_n \cap V \tag{6.186}$$

contient l'intersection des B_n. Par le théorème 6.136 des fermés emboités (que nous utilisons parce que E est métrique et complet), cette intersection est non vide.

Ouvert d'un espace de Baire

□

Une des application du théorème de Baire est le théorème de Banach-Steinhaus 24.8.

Chapitre 7

Espaces vectoriels

7.1 Parties libres, génératrices, bases et dimension

Définition 7.1 (Partie libre)**.**
Si E est un espace vectoriel, une partie A de E est ***libre*** *si pour tout choix d'un nombre fini d'éléments $\{u_i\}_{i=1,\dots,n}$, l'égalité*

$$a_1u_1 + \cdots + a_nu_n = 0 \tag{7.1}$$

implique $a_i = 0$ pour tout i (ici les a_i sont dans le corps de base).

Une partie infinie est libre si toute ses parties finies le sont.

Remarque 7.2.
Notons que le vecteur nul n'est dans aucune partie libre, ne fut-ce que parce que $a0 = 0$ n'implique pas $a = 0$.

Si A est une partie de l'espace vectoriel E nous notons $\text{Span}(A)$ l'ensemble des combinaisons linéaires finies d'éléments de A. Les coefficients de ces combinaisons linéaires sont dans le corps de base $\mathbb{K}$.

Définition 7.3 (Partie génératrice)**.**
Une partie B d'un espace vectoriel E est ***génératrice*** *si* $\text{Span}(B) = E$.

Remarque 7.4.
Ces définitions demandent des commentaires en dimension infinie [1].

(1) Tout élément peut être écrit comme combinaison linéaire finie d'une partie génératrice. Cela ne signifie pas que nous pouvons extraire une partie finie qui convient pour tous les éléments à la fois. Lorsque l'espace est de dimension infinie, ceci est particulièrement important.

(2) La définition séparée de liberté dans le cas des parties infinies a son importance lorsqu'on parle d'espaces vectoriels de dimension infinies (en dimension finie, aucune partie infinie n'est libre) parce que cela fera une différence entre une base algébrique et une base hilbertienne par exemple.

Définition 7.5 (Base)**.**
Une ***base*** *de l'espace vectoriel E est une partie à la fois génératrice et libre.*

Proposition 7.6 ([1])**.**
Tout élément non nul d'un espace vectoriel se décompose de façon unique en combinaison linéaire finie d'éléments d'une base.

Démonstration. Soit un espace vectoriel E et une base $\{e_i\}_{i\in I}$ où I est un ensemble a priori quelconque. Soit $v \in E$. Vu que $E = \text{Span}\{e_i\}_{i\in I}$, il existe une partie finie J de I et des coefficients

1. Nous n'avons pas encore définit le concept de dimension, mais nous nous adressons au lecteur trop pressé.

$\{v_j\}_{j\in J}$ dans $\mathbb{K}$ tels que

$$v = \sum_{j\in J} v_j e_j. \tag{7.2}$$

Cela donne l'existence.

En ce qui concerne l'unicité, soient J et K des parties finies de I et des coefficients $\{v_j\}_{j\in J}$ et $\{w_k\}_{k\in K}$ tels que

$$v = \sum_{j\in J} v_j e_j = \sum_{k\in K} v_k e_k. \tag{7.3}$$

Nous posons $L = J \cup K$ et, pour $l \in L$,

$$\alpha_l = \begin{cases} v_l & \text{si } l \in J\backslash K \\ w_l & \text{si } l \in K\backslash J \\ v_l - w_l & \text{si } l \in K \cap J. \end{cases} \tag{7.4}$$

Nous avons alors

$$\sum_{l\in L} \alpha_l e_l = 0, \tag{7.5}$$

ce qui implique que $\alpha_l = 0$ pour tout $l \in L$ parce que la partie $\{e_i\}_{i\in I}$ est libre et que L est finie.

L'unicité de la décomposition de v signifie que

$$\{j \in J \text{ tel que } v_j \neq 0\} = \{k \in K \text{ tel que } w_k \neq 0\} \tag{7.6}$$

et que pour l dans cet ensemble, $v_l = w_l$.

Soit $j \in J$; il y a deux possibilités : $j \in J\backslash K$ et $j \in J \cap K$. Dans le premier cas nous avons déjà vu que $\alpha_j = v_j = 0$. Dans le second cas, $\alpha_j = v_j - w_j = 0$, c'est à dire $v_j = w_j$.

Donc $j \in J$ vérifiant $v_j \neq 0$ implique $j \in J \cap K$ et l'égalité des coefficients. Idem avec $k \in K$ tel que $w_k \neq 0$ implique $k \in J \cap K$. □

Définition 7.7.
*Un espace vectoriel est **de type fini** s'il contient une partie génératrice finie.*

Nous verrons dans les résultats qui suivent que cette définition est en réalité inutile parce qu'une espace vectoriel sera de type fini si et seulement s'il est de dimension finie.

Lemme 7.8.
Si E a une famille génératrice de cardinal n, alors toute famille de $n+1$ éléments est liée.

Démonstration. Nous procédons par récurrence sur n – qui n'est pas exactement la dimension d'un espace vectoriel fixé. Pour $n = 1$, nous avons $E = \langle e\rangle$ et donc si $v_1, v_2 \in E$ nous avons $v_1 = \lambda_1 e$, $v_2 = \lambda_2 e$ et donc $\lambda_2 v_1 - \lambda_1 v1 = 0$. Cela prouve le résultat dans le cas de la dimension 1.

Supposons maintenant que le résultat soit vrai pour $k < n$, c'est à dire que pour tout espace vectoriel contenant une partie génératrice de cardinal $k < n$, les parties de $k+1$ éléments sont liées. Soit maintenant un espace vectoriel muni d'une partie génératrice $G = \{e_1, \dots, e_n\}$ de n éléments, et montrons que toute partie $V = \{v_1, \dots, v_{n+1}\}$ contenant $n+1$ éléments est liée. Dans nos notations nous supposons que les e_i sont des vecteurs distincts et les v_i également. Nous les supposons également tous non nuls. Étant donné que $\{e_i\}$ est génératrice nous pouvons définir les nombres λ_{ij} par

$$v_i = \sum_{k=1}^{n} \lambda_{ij} e_j \tag{7.7}$$

Vu que

$$v_{n+1} = \sum_{k=1}^{n} \lambda_{n+1,k} e_k \neq 0, \tag{7.8}$$

quitte à changer la numérotation des e_i nous pouvons supposer que $\lambda_{n+1,n} \neq 0$. Nous considérons les vecteurs

$$w_i = \lambda_{n+1,n} v_i - \lambda_{i,n} v_{n+1}. \tag{7.9}$$

En calculant un peu,

$$w_i = \lambda_{n+1,n} \sum_k \lambda_{i,k} e_k - \lambda_{i,n} \sum_k \lambda_{n+1,k} e_k \tag{7.10a}$$

$$= \sum_{k=1}^{n-1} \big(\lambda_{n+1,n}\lambda_{i,k} - \lambda_{i,n}\lambda_{n+1,}\big) e_k \tag{7.10b}$$

parce que les termes en e_n se sont simplifiés. Donc la famille $\{w_1, \dots, w_n\}$ est une famille de n vecteurs dans l'espace vectoriel $\text{Span}\{e_1, \dots, e_{n-1}\}$; elle est donc liée par l'hypothèse de récurrence. Il existe donc des nombres $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ non tous nuls tels que

$$0 = \sum_{i=1}^{n} \alpha_i w_i = \sum_{i=1}^{n} \alpha_i \lambda_{n+1,n} v_i - \left(\sum_{i=1}^{n} \alpha_i \lambda_{i,n} \right) v_{n+1}. \tag{7.11}$$

Vu que $\lambda_{n+1,n} \neq 0$ et que parmi les α_i au moins un est non nul, nous avons au moins un des produits $\alpha_i \lambda_{n+1,n}$ qui est non nul. Par conséquent (7.11) est une combinaison linéaire nulle non triviale des vecteurs de $\{v_1, \dots, v_{n+1}\}$. Cette partie est donc liée. □

Lemme 7.9.
Soit L une partie libre et G une partie génératrice. Soit B une partie maximale parmi les parties libres L' telles que $L \subset L' \subset G$. Alors B est une base.

Qu'entend-on par «maximale» ? La partie B doit être libre, contenir L, être contenue dans G et de plus avoir la propriété que $\forall x \in G \backslash B$, la partie $B \cup \{x\}$ est liée.

Démonstration. D'abord si G est une base, alors toutes les parties de G sont libres et le maximum est $B = G$. Dans ce cas le résultat est évident. Nous supposons donc que G est liée.

La partie $B = \{b_1, \dots, b_l\}$ est libre parce qu'on l'a prise parmi les libres. Montrons que B est génératrice. Soit $x \in G \backslash B$; par hypothèse de maximalité, $B \cup \{x\}$ est liée, c'est à dire qu'il existe des nombres λ_i, λ_x non tous nuls tels que

$$\sum_{i=1}^{l} \lambda_i b_i + \lambda_x x = 0. \tag{7.12}$$

Si $\lambda_x = 0$ alors un de λ_i doit être non nul et l'équation (7.12) devient une combinaison linéaire nulle non triviale des b_i, ce qui est impossible parce que B est libre. Donc $\lambda_x \neq 0$ et

$$x = \frac{1}{\lambda_x} \sum_{i=1}^{l} \lambda_i b_i. \tag{7.13}$$

Donc tous les éléments de $G \backslash B$ sont des combinaisons linéaires des éléments de B, et par conséquent, G étant génératrice, tous les éléments de E sont combinaisons linéaires d'éléments de B. □

Théorème 7.10 (Théorème de la base incomplète)**.**
Soit E un espace vectoriel de type fini sur le corps $\mathbb{K}$.

(1) Si L est une partie libre et si G est une partie génératrice contenant L, alors il existe une base B telle que $L \subset B \subset G$.

(2) Toute partie libre peut être étendue en une base.

(3) Toutes les bases sont finies et ont même cardinal.

Démonstration. Point par point.

(1) Vu que E est de type fini, il admet une partie génératrice G de cardinal fini n. Donc une partie libre est de cardinal au plus n par le lemme 7.8. Soit L, une partie libre contenue dans G (ça existe : par exemple $L = \varnothing$). La partie B maximalement libre contenue dans G et contenant L est une base par le lemme 7.9.

(2) Notons que puisque E lui-même est générateur, le point (1) implique que toute partie libre peut être étendue en une base.

(3) Soient B et B', deux bases. En particulier B est génératrice et B' est libre, donc le lemme 7.8 indique que $\text{Card}(B') \leqslant \text{Card}(B)$. Par symétrie on a l'inégalité inverse. Donc $\text{Card}(B) = \text{Card}(B')$.

□

Remarque 7.11.
Le théorème de la base incomplete 7.10(2) est ce qui permet de construire une base d'une espace vectoriel en « commençant par» une base d'un sous-espace. En effet si H est un sous-espace de E alors une base de H est une partie libre de E et donc peut être étendue en une base de E.

Définition 7.12.
*La **dimension** d'un espace vectoriel de type fini est la cardinal d'une de ses bases.*

Exemple 7.13
Il existe une infinité de bases de $\mathbb{R}^m$. On peut démontrer que le cardinal de toute base de $\mathbb{R}^m$ est m, c'est à dire que toute base de $\mathbb{R}^m$ possède exactement m éléments.

La base de $\mathbb{R}^m$ qu'on dit **canonique** (c.à.d. celle qu'on utilise tout le temps) est $\mathcal{B} = \{e_1, \dots, e_m\}$, où le vecteur e_j est

$$e_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{j-ème} \quad .$$

La composante numéro j de e_i est 1 si $i = j$ et 0 si $i \neq j$. Cela s'écrit $(e_i)_j = \delta_{ij}$ où δ est le **symbole de Kronecker** défini par

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \tag{7.14}$$

Les éléments de la base canonique de $\mathbb{R}^m$ peuvent donc être écrits $e_i = \sum_{k=1}^{m} \delta_{ik} e_k$. △

Le théorème suivant est essentiellement une reformulation du théorème 7.10.

Théorème 7.14.
Soit E un espace vectoriel de dimension finie et $\{e_i\}_{i\in I}$ une partie génératrice de E.

(1) Il existe $J \subset I$ tel que $\{e_i\}_{i\in J}$ est une base. Autrement dit : de toute partie génératrice nous pouvons extraire une base.

(2) Soit $\{f_1, \dots, f_l\}$ une partie libre. Alors nous pouvons la compléter en utilisant des éléments e_i. C'est à dire qu'il existe $J \subset I$ tel que $\{f_k\} \cup \{e_i\}_{i\in J}$ soit une base.

(3) Si E est un espace vectoriel de dimension n, alors toute partie libre contenant n éléments est une base.

Démonstration. Nous ne prouvons que le troisième point. Une partie libre contenant n éléments peut être étendue en une base ; si ladite extension est non triviale (c'est à dire qu'on ajoute vraiment

au moins un élément) une telle base contiendra une partie de $n+1$ éléments qui serait liée par le lemme 7.8. □

Soit F un sous-espace vectoriel de l'espace vectoriel E. La **codimension** de F dans E est

$$\operatorname{codim}_E(F) = \dim(E/F). \tag{7.15}$$

7.2 Applications linéaires

7.2.1 Définition

Définition 7.15.
*Soient des espaces vectoriels E et F sur le corps $\mathbb{K}$. Une application $T\colon E \to F$ est dite **linéaire** si*

- *$T(x+y) = T(x) + T(y)$ pour tout x et y dans E,*
- *$T(\lambda x) = \lambda T(x)$ pour tout λ dans $\mathbb{K}$ et x dans E.*

Si vous avez bien suivi, les égalités dans la définition 7.15 sont des égalités dans F.

Exemple 7.16
Pour tout b dans $\mathbb{R}$ la fonction $T_b(x) = bx$ est une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$. En effet,

- $T_b(x+y) = b(x+y) = bx + by = T_b(x) + T_b(y)$,
- $T_b(ax) = b(ax) = abx = aT_b(x)$.

De la même façon on peut montrer que la fonction T_λ définie par $T_\lambda(x) = bx$ est un application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^m$ pour tout λ dans $\mathbb{R}$ et m dans $\mathbb{N}$. △

Définition 7.17 (Quelque ensembles d'applications linéaires)**.**
Soient E et F des espaces vectoriels.

- *L'ensemble des applications linéaires de E vers F est noté $\mathcal{L}(E,F)$.*
- *Une application linéaire $E \to E$ est un **endomorphisme** de E. L'ensemble des endomorphisme de E est noté* $\operatorname{End}(E)$.
- *Un endomorphisme bijectif est un **automorphisme**. L'ensemble des automorphismes de E est noté* $\operatorname{Aut}(E)$.
- *Une application linéaire bijective $E \to F$ est un **isomorphisme** d'espace vectoriel. L'ensemble des isomorphismes $E \to F$ est noté* $\operatorname{GL}(E,F)$.

Remarque 7.18.
Les ensembles définis en 7.17 concernent la structure d'espace vectoriel seulement. Lorsque nous verrons la notion d'espace vectoriel normé,nous demanderons de plus la continuité, laquelle n'est pas automatique en dimension infinie. Voir aussi les définitions 10.19.

Définition 7.19.
Si V et W sont des espaces vectoriels nous munissons $\mathcal{L}(V,W)$ d'une structure d'espace vectoriel en définissant la somme et le produit par un scalaire de la façon suivante. Si T et U sont des élément de $\mathcal{L}(V,W)$ et si λ est un réel, nous définissons les éléments $T+U$ et λT par

(1) $(T+U)(x) = T(x) + U(x)$;

(2) $(\lambda T)(x) = \lambda T(x)$

pour tout $x \in V$.

7.20.
Pour bien faire, il faudrait vérifier que la définition 7.19 produit effectivement une structure d'espace vectoriel sur $\mathcal{L}(V,W)$.

Proposition 7.21.
Si E et F sont des espaces vectoriels de dimension n et si $\{e_i\}_{i=1,\dots,n}$ et $\{f_i\}_{i=1,\dots,n}$ sont des bases

respectivement de F et F, alors il existe une unique application linéaire $A\colon E \to F$ telle que $E(e_i) = f_i$ pour tout i.

Démonstration. En deux parties.

Existence Soit $v \in E$. Vu que $\{e_i\}$ est une base, il se décompose de façon unique en $v = \sum_i v_i e_i$. Alors définir

$$A(v) = \sum_i v_i f_i \tag{7.16}$$

est une bonne définition et satisfait aux exigences.

Unicité Soient A et B satisfaisant aux exigences. Alors pour tout i nous avons $A(e_i) = B(e_i)$. Si $v \in E$ s'écrit de la forme $v = \sum_i v_i e_i$ alors la linéarité impose $A(v) = \sum_i v_i A(e_i) = \sum_i v_i B(e_i) = B(v)$. Donc $A = B$.

□

7.2.2 Rang

Définition 7.22.
*Si $T\colon V \to W$ est une application linéaire, son **rang** est la dimension de son image.*

Le théorème suivant est valable également en dimension infinie ; ce sera une des rares incursions en dimension infinie de ce chapitre.

Théorème 7.23 (Théorème du rang)**.**
Soient E et F deux espaces vectoriels (de dimensions finies ou non) et soit $f\colon E \to F$ une application linéaire. Alors le rang[2] *de f est égal à la codimension du noyau, c'est à dire*

$$\operatorname{rang}(f) + \dim \ker f = \dim E. \tag{7.17}$$

Dans le cas de dimension infinie afin d'éviter les problèmes d'arithmétique avec l'infini nous énonçons le théorème en disant que si $(u_s)_{s\in S}$ est une base de $\ker(f)$ et si $\big(f(v_t)\big)_{t\in T}$ est une base de $\operatorname{Image}(f)$ alors $(u_s)_{s\in s} \cup (v_t)_{t\in T}$ est une base de E.

Démonstration. Nous devons montrer que

$$(u_s)_{s\in S} \cup (v_t)_{t\in T} \tag{7.18}$$

est libre et générateur.

Soit $x \in E$. Nous définissons les nombres x_t par la décomposition de $f(x)$ dans la base $\big(f(v_t)\big)$:

$$f(x) = \sum_{t\in T} x_t f(v_t). \tag{7.19}$$

Ensuite le vecteur $x = \sum_t x_t v_t$ est dans le noyau de f, par conséquent nous le décomposons dans la base (u_s) :

$$x - \sum_t x_t v_t = \sum_s \in S x_s u_s. \tag{7.20}$$

Par conséquent

$$x = \sum_s x_s u_s + \sum_t x_t v_t. \tag{7.21}$$

En ce qui concerne la liberté nous écrivons

$$\sum_t x_t v_t + \sum_s x_s u_s = 0. \tag{7.22}$$

2. Définition 7.22.

En appliquant f nous trouvons que

$$\sum_t x_t f(v_t) = 0 \tag{7.23}$$

et donc que les x_t doivent être nuls. Nous restons avec $\sum_s x_s u_s = 0$ qui à son tour implique que $x_s = 0$. □

Un exemple d'utilisation de ce théorème en dimension infinie sera donné dans le cadre du théorème de Fréchet-Riesz, théorème 23.18.

Corollaire 7.24.
Si E est un espace vectoriel de dimension finie, alors un endomorphisme de E est bijectif si et seulement si il est injectif si et seulement si il est surjectif.

Démonstration. Si un endomorphisme $f\colon E \to E$ est surjectif, alors $\text{rang}(f) = \dim(E)$, ce qui donne, par le théorème du rang 7.23, $\dim\big(\ker(f)\big) = 0$, c'est à dire que f est injectif.

De la même façon, si f est injective, alors $\dim\big(\ker(f)\big) = 0$, ce qui donne $\text{rang}(f) = \dim(E)$ ou encore que f est surjective. □

Une conséquence du théorème du rang est que les endomorphismes ont un inverse à gauche et à droite égaux (lorsqu'ils existent).

Corollaire 7.25.
Soit un endomorphisme f d'un espace vectoriel de dimension finie. Si f admet un inverse à gauche, alors

(1) f est bijective,

(2) f admet également un inverse à droite,

(3) ils sont égaux.

Tout cela tient également en remplaçant «gauche» par «droite».

Démonstration. Soit g, un inverse à gauche de f : $gf = \text{Id}$. Cela implique que f est injective et que g est surjective, et donc qu'elles sont toutes deux bijectives par le corollaire 7.24. Vu que f est bijective, elle admet également un inverse à droite, soit h. Nous avons : $gf = \text{Id}$ et $fh = \text{Id}$.

Alors $gfh = h$ parce que $gf = \text{Id}$, mais également $gfh = g$ parce que $fh = \text{Id}$. Donc $g = h$. □

C'est ce corollaire qui nous permet d'écrire f^{-1} sans plus de précisions dès que f est une bijection.

Exemple 7.26(Ne fonctionne pas en dimension infinie)
Tout cela ne fonctionne pas en dimension infinie. Par exemple avec une base $\{e_k\}_{k\in\mathbb{N}}$ nous pouvons considérer l'opérateur

$$f(e_k) = e_{k+1}. \tag{7.24}$$

Il est injectif, mais pas surjectif. Si on pose

$$g(e_k) = \begin{cases} e_{k-1} & \text{si } k \geqslant 1 \\ 0 & \text{si } k = 0 \end{cases} \tag{7.25}$$

alors nous avons $gf = \text{Id}$, mais pas $fg = \text{Id}$ parce que ce $(fg)(e_0) = 0$. △

7.2.3 Matrices équivalentes et similaires

Définition 7.27.
Deux relations d'équivalence entre les matrices.

*(1) Deux matrices A et B sont **équivalentes** dans $\mathbb{M}(n,\mathbb{K})$ s'il existe $P, Q \in \text{GL}(n,\mathbb{K})$ telles que $A = PBQ^{-1}$.*

(2) Deux matrices sont ***semblables*** *s'il existe une matrice* $P \in \mathrm{GL}(n,\mathbb{K})$ *telle que* $A = PBP^{-1}$.

Lemme 7.28.
Une matrice de rang r *dans* $\mathbb{M}(n,\mathbb{K})$ *est équivalente à la matrice par blocs*

$$J_r = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{7.26}$$

Démonstration. Nous devons prouver que pour toute matrice $A \in \mathbb{M}(n,\mathbb{K})$ de rang r, il existe $P, Q \in \mathrm{GL}(n,\mathbb{K})$ telles que $QAP = J_r$. Soit $\{e_i\}$ la base canonique de $\mathbb{K}^n$, puis $\{f_i\}$ une base telle que $Af_i = 0$ dès que $i > r$.

Nous considérons la matrice inversible P telle que $Pe_i = f_i$. Elle vérifie

$$APe_i = Af_i = \begin{cases} 0 & \text{si } i > r \\ \neq 0 & \text{sinon.} \end{cases} \tag{7.27}$$

La matrice AP se présente donc sous la forme

$$AP = \begin{pmatrix} M & 0 \\ * & 0 \end{pmatrix} \tag{7.28}$$

où M est une matrice $r \times r$. Nous considérons maintenant une base $\{g_i\}_{i=1,\dots,n}$ dont les r premiers éléments sont les r premières colonnes de AP et une matrice inversible Q telle que $Qg_i = e_i$. Alors

$$QAPe_i = \begin{cases} e_i & \text{si } i < r \\ 0 & \text{sinon.} \end{cases}. \tag{7.29}$$

Cela signifie que QAP est la matrice J_r. □

Corollaire 7.29 (Équivalence et rang)**.**
Deux matrices sont équivalentes[3] *si et seulement si elles ont même rang.*

Démonstration. D'abord il y a des implicites dans l'énoncé. Vu que nous voulons soit pas hypothèse soit par conclusion que les matrices A et B soient équivalentes, nous supposons qu'elles ont même dimension. Soient donc A et B deux matrices carré représentant des applications linéaires de l'espace vectoriel $E = \mathbb{K}^n$ vers lui-même.

Par le lemme 7.28, deux matrices de même rang r sont équivalentes à J_r. Elles sont donc équivalentes entre elles.

Inversement, supposons que A et B soient deux matrices équivalentes : $A = PBQ^{-1}$ avec P et B inversibles. Alors

$$\begin{aligned} \mathrm{Image}(PBQ^{-1}) &= \{PBQ^{-1}v \text{ tel que } v \in E\} && \text{(7.30a)} \\ &= PB\underbrace{\{Q^{-1}v \text{ tel que } v \in E\}}_{=E} && \text{(7.30b)} \\ &= P\big(B(E)\big). && \text{(7.30c)} \end{aligned}$$

L'ensemble $B(E)$ est un sous-espace vectoriel de E. Vu que le rang de P est maximum, la dimension de $P\big(B(E)\big)$ est la même que celle de $B(E)$. Par conséquent

$$\dim\Big(\mathrm{Image}(PBQ^{-1})\Big) = \dim\big(B(E)\big) = \mathrm{rang}(B). \tag{7.31}$$

Le membre de gauche de cela n'est autre que $\mathrm{rang}(A) = \dim\big(\mathrm{Image}(PBQ^{-1})\big)$. □

3. Définition 7.27(1).

7.3 Matrice d'une application linéaire

7.3.1 Écriture dans une base

Exemple 7.30
Soit $m = n$. On fixe λ dans $\mathbb{R}$ et v dans $\mathbb{R}^m$. L'application U_λ de $\mathbb{R}^m$ dans $\mathbb{R}^m$ définie par $U_\lambda(x) = \lambda x + v$ n'est pas une application linéaire, parce que

$$U_\lambda(ax) = \lambda(ax) + v \neq \lambda(bx + v) = aU_\lambda(x).$$

△

Exemple 7.31
Soit A une matrice fixée de $\mathcal{M}_{n\times m}$. La fonction $T_A\colon \mathbb{R}^m \to \mathbb{R}^n$ définie par $T_A(x) = Ax$ est une application linéaire. En effet,
— $T_A(x + y) = A(x + y) = Ax + Ay = T_A(x) + T_A(y)$,
— $T_A(ax) = A(ax) = a(Ax) = aT_A(x)$.

△

On peut observer que, si on identifie $\mathcal{M}_{1\times 1}$ et $\mathbb{R}$, on obtient le premier exemple comme cas particulier.

Proposition 7.32.
Toute application linéaire T de $\mathbb{R}^m$ dans $\mathbb{R}^n$ s'écrit de manière unique par rapport aux bases canoniques de $\mathbb{R}^m$ et $\mathbb{R}^n$ sous la forme

$$T(x) = Ax,$$

avec A dans $\mathcal{M}_{n\times m}$.

Démonstration. Soit x un vecteur dans $\mathbb{R}^m$. On peut écrire x sous la forme $x = \sum_{i=1}^m x_i e_i$. Comme T est une application linéaire on a

$$T(x) = \sum_{i=1}^m x_i T(e_i).$$

Les images de la base de $\mathbb{R}^m$, $T(e_j)$, $j = 1, \ldots, m$, sont des éléments de $\mathbb{R}^n$, donc on peut les écrire sous la forme de vecteurs

$$T(e_i) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}.$$

On obtient alors

$$T(x) = \sum_{i=1}^m x_i T(e_i) = \sum_{i=1}^m x_i \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix} = \begin{pmatrix} a_{11} \ldots a_{1m} \\ \vdots \ddots \vdots \\ a_{n1} \ldots a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = Ax.$$

□

Définition 7.33.
*Une application $S : \mathbb{R}^m \to \mathbb{R}^n$ est dite **affine** si elle est la somme d'une application linéaire et d'une application constante. Autrement dit, S est affine s'il existe $T : \mathbb{R}^m \to \mathbb{R}^n$, linéaire, telle que $S(x) - T(x)$ soit un vecteur constant dans $\mathbb{R}^n$.*

Exemple 7.34
Les exemples les plus courants d'applications affines sont les droites et les plans ne passant pas par l'origine.

Les droites Une droite dans $\mathbb{R}^2$ (ou $\mathbb{R}^3$) qui ne passe pas par l'origine est le graphe d'une fonction de la forme $s(x) = ax + b$ (ou $s(t) = ux + v$, avec u et v dans $\mathbb{R}^2$). On reconnait ici la fonction de l'exemple 7.30.

Les plans De la même façon nous savons que tout plan qui ne passe pas par l'origine dans $\mathbb{R}^3$ est le graphe d'une application affine, $P(x, y) = (a, b)^T \cdot (x, y)^T + (c, d)^T$.

△

Lemme 7.35 ([83])**.**
Soit une application linéaire $f\colon E \to F$.

(1) L'application f est injective si et seulement s'il existe $g\colon F \to E$ telle que $g \circ f = \mathrm{Id}\,|_E$.

(2) L'application f est surjective si et seulement s'il existe $g\colon F \to E$ telle que $f \circ g = \mathrm{Id}\,|_F$.

Démonstration. Nous démontrons séparément les deux affirmations.

(1) Si f est injective, alors $f\colon E \to \mathrm{Image}(f)$ est un isomorphisme. Si V est un supplémentaire de $\mathrm{Image}(f)$ dans F (c'est à dire $F = \mathrm{Image}(f) \oplus V$) alors nous pouvons poser $g(x + v) = f^{-1}(x)$ où $x + v$ est la décomposition (unique) d'un élément de F en $x \in \mathrm{Image}(f)$ et $v \in V$. Avec cela nous avons bien $g \circ f = \mathrm{Id}$.

Inversement, s'il existe $g\colon F \to E$ telle que $g \circ f = \mathrm{Id}$ alors $f\colon E \to E$ doit être injective. Parce que si $f(x) = 0$ avec $x \neq 0$ alors $(g \circ f)(x) = 0 \neq x$.

(2) Si f est surjective nous pouvons choisir des éléments $x_1, \dots, x_p$ dans E tels que $\{f(x_i)\}$ soit une base de F. Ensuite nous définissons

$$\begin{aligned} g\colon F &\to E \\ \sum_k a_k f(x_k) &\mapsto \sum_k x_k. \end{aligned} \tag{7.32}$$

Cela donne $f \circ g = \mathrm{Id}\,|_F$ parce que si $v \in F$ alors $v = \sum_k v_k f(x_k)$ avec $v_k \in \mathbb{K}$, et nous avons

$$(f \circ g)(v) = \sum_k v_k (f \circ g) f(x_k) = f\left(\sum_k v_k x_k\right) = \sum_k f(x_k) = v. \tag{7.33}$$

Inversement, s'il existe $g\colon F \to E$ tel que $f \circ g = \mathrm{Id}$ alors f soit être surjective parce que

$$F = \mathrm{Image}(f \circ g) = f\big(\mathrm{Image}(g)\big) \subset \mathrm{Image}(f). \tag{7.34}$$

□

7.3.2 Changement de base : vecteurs de base

Soit un espace vectoriel V muni d'une base $\{e_i\}$ et une application linéaire inversible $Q\colon V \to V$ dont nous notons (Q_{ij}) les coordonnées dans la base $\{e_i\}$. Nous considérons aussi la base $e'_i = Qe_i$.

Nous exprimons les vecteurs e'_i en termes des vecteurs e_i et inversement de la façon suivante. D'abord

$$e'_i = Qe_i = \sum_{kl} Q_{kl} (e_i)_k e_l = \sum_l Q_{il} e_l \tag{7.35}$$

donc

$$e'_i = \sum_k Q_{ik} e_k. \tag{7.36}$$

7.36.
Afin de ne pas se tromper dans les indices, il est bon d'utiliser systématiquement les noms "$ijk\dots$" pour les indices de la base $\{e_i\}$ et les noms "$\alpha\beta\gamma\cdots$" pour les indices de la base $\{e'_\alpha\}$. Nous écrivons donc

$$e'_\alpha = \sum_i Q_{i\alpha} e_i. \tag{7.37}$$

Pour trouver la transformation inverse, nous multiplions par $(Q^{-1})_{\alpha j}$ et nous sommons sur α :

$$\sum_{\alpha}(Q^{-1})_{\alpha j}e'_{\alpha} = \sum_{i\alpha} Q_{i\alpha}Q^{-1}_{\alpha j}e_i = e_j \tag{7.38}$$

donc

$$e_j = \sum_{\alpha}(Q^{-1})_{\alpha j}e'_{\alpha}. \tag{7.39}$$

Quelque remarques :

(1) Notons que nous ne pouvons pas écrire formellement quelque chose comme $f = Qe$.

(2) La matrice Q vient toujours avec les indices de type $Q_{i\alpha}$ (latin en premier et grec en second) tandis que la matrice de l'application inverse Q^{-1} vient toujours avec les indices de type $(Q^{-1})_{\alpha i}$.
Ceci est un puissant outil pour ne pas se tromper lors des manipulations d'indices et en particulier lorsque nous renommons des indices.

7.3.3 Changement de base : coordonnées

Un vecteur x peut s'écrire des façons suivantes dans les deux bases considérées :

$$x = \sum_i x_i e_i = \sum_{\alpha} x'_{\alpha}e'_{\alpha}. \tag{7.40}$$

En remplaçant e_i par sa valeur (7.39) nous obtenons l'égalité

$$\sum_{k\alpha} x_k(Q^{-1})_{\alpha k}e'_{\alpha} = \sum_{\alpha} x'_{\alpha}e'_{\alpha} \tag{7.41}$$

et par unicité de la décomposition d'un vecteur dans une base nous déduisons

$$x'_{\alpha} = \sum_k (Q^{-1})_{\alpha k}x_k. \tag{7.42}$$

De même en replaçant e'_{α} par sa valeur (7.37) nous trouvons

$$x_i = \sum_{\alpha} Q_{i\alpha}x'_{\alpha}. \tag{7.43}$$

Attention : il n'est pas question d'écrire quelque chose comme $x = Qx'$ parce qu'il n'y a pas de vecteur x'.

7.3.4 Changement de base : matrice d'une application linéaire

Soit une application linéaire $f\colon V \to V$ de matrice A dans la base $\{e_i\}$ et de matrice A' dans la base $\{e'_{\alpha}\}$. Notre but est maintenant d'exprimer A' en termes de A et de Q.

Nous avons l'égalité

$$f(x) = \sum_{kl} A_{kl}x_l e_k = \sum_{\alpha\beta} A'_{\alpha\beta}x'_{\beta}e'_{\alpha}. \tag{7.44}$$

Il suffit d'y substituer les valeurs (7.43) et (7.39) :

$$f(x) = \sum_{kl} A_{kl}\sum_{\beta} Q_{l\beta}x'_{\beta}\sum_{\alpha}(Q^{-1})_{\alpha k}e'_{\alpha} = \sum_{\alpha\beta}(Q^{-1}AQ)_{\alpha\beta}x'_{\beta}e'_{\alpha}. \tag{7.45}$$

Cela est égal à $\sum_{\alpha\beta} A'_{\alpha\beta}x'_{\beta}e'_{\alpha}$ pour tout x. En tenant encore compte de l'unicité de la décomposition du vecteur $f(x)$ dans la base $\{e'_{\alpha}\}$ nous déduisons

$$A' = Q^{-1}AQ. \tag{7.46}$$

Remarque 7.37.
L'égalité (7.46) est une égalité de matrices, et non une égalité d'applications linéaire. Il n'est pas question d'écrire quelque chose comme $f' = Q^{-1} \circ f \circ Q$.

7.3.5 Changement de base : matrice d'une forme bilinéaire

Soit une forme bilinéaire $q\colon V \times V \to \mathbb{R}$ de matrice (q_{ij}) dans la base $\{e_i\}$ et de matrice $q'_{\alpha\beta}$ dans la base $\{e'_\alpha\}$. Nous notons le changement de variables $e'_\alpha = \sum_i Q_{i\alpha} e_i$. Nous avons l'égalité

$$q(x,y) = \sum_{ij} q_{ij} x_i y_j = \sum_{\alpha\beta} q'_{\alpha\beta} x'_\alpha y'_\beta. \tag{7.47}$$

En remplaçant x_i et y_j par leurs valeurs en termes des x'_α et y'_β nous trouvons

$$q(x,y) = \sum_{ij\alpha\beta} q_{ij} Q_{i\alpha} x'_\alpha Q_{j\beta} y'_\beta, \tag{7.48}$$

c'est à dire

$$q' = Q^t q Q. \tag{7.49}$$

Encore une fois, cette égalité est une égalité de matrices, et non de formes bilinéaires, et encore moins d'applications linéaires.

Remarque 7.38.
La «loi de transformation» (7.49) n'est pas la même que la loi de transformation (7.46).

7.4 Dualité

Proposition 7.39.
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 carré de déterminant non nul contenue dans A.

Définition 7.40.
*Si E est un espace vectoriel sur $\mathbb{K}$, le **dual algébrique** de E est l'ensemble des formes linéaires sur E. Le **dual topologique** est l'ensemble des formes linéaires continues de E vers $\mathbb{K}$.*

En dimension infinie, ces deux notions ne coïncident pas, voir la proposition 10.18 et la remarque 10.22.

7.4.1 Orthogonal

Soit E, un espace vectoriel, et F une sous-espace de E. L'**orthogonal** de F est la partie $F^\perp \subset E^*$ donnée par

$$F^\perp = \{\alpha \in E^* \text{ tel que } \forall x \in F, \alpha(x) = 0\}. \tag{7.50}$$

Cette définition d'orthogonal via le dual n'est pas du pur snobisme. En effet, la définition «usuelle» qui ne parle pas de dual,

$$F^\perp = \{y \in E \text{ tel que } \forall x \in F, y \cdot x = 0\}, \tag{7.51}$$

demande la donnée d'un produit scalaire. Évidemment dans le cas de $\mathbb{R}^n$ munie du produit scalaire usuel et de l'identification usuelle entre $\mathbb{R}^n$ et $(\mathbb{R}^n)^*$ via une base, les deux notions d'orthogonal coïncident.

Si $B \subset E^*$, on note B^o son orthogonal :

$$B^o = \{x \in E \text{ tel que } \omega(x) = 0\, \forall \omega \in B\}. \tag{7.52}$$

Notons qu'on le note B^o et non $B^\perp$ parce qu'on veut un peu s'abstraire du fait que $(E^*)^* = E$. Du coup on impose que B soit dans un dual et on prend une notation précise pour dire qu'on remonte au pré-dual et non qu'on va au dual du dual.

La définition (7.50) est intrinsèque : elle ne dépend que de la structure d'espace vectoriel.

Proposition 7.41.
Soit E un espace vectoriel, et F un sous-espace vectoriel. Alors nous avons

$$\dim F + \dim F^{\perp} = \dim E. \tag{7.53}$$

Démonstration. Soit $\{e_1, \dots, e_p\}$ une base de F que nous complétons en une base $\{e_1, \dots, e_n\}$ de E par le théorème 7.10. Soit $\{e_1^*, \dots, e_n^*\}$ la base duale. Alors nous prouvons que $\{e_{p+1}^*, \dots, e_n^*\}$ est une base de $F^{\perp}$.

Déjà c'est une partie libre en tant que partie d'une base.

Ensuite ce sont des éléments de F^{perp} parce que si $i \leqslant p$ et si $k \geqslant 1$ nous avons $e_{p+k}^*(e_i^*) = 0$; donc oui, $e_{p+k}^* \in F^{\perp}$.

Enfin $F^{\perp} \subset \text{Span}\{e_{p+1}^*, e_n^*\}$ parce que si $\omega = \sum_{k=1}^{n} \omega_k e_k^*$, alors $\omega(e_i) = \omega_i$, mais nous savons que si $\omega \in F^{\perp}$, alors $\omega(e_i) = 0$ pour $i \leqslant p$. Donc $\omega = \sum_{i=p+1}^{n} \omega_k e_k^*$. □

7.4.2 Transposée

Définition 7.42.
*Si $f : E \to F$ est une application linéaire entre deux espaces vectoriels, la **transposée** est l'application $f^t : F^* \to E^*$ donnée par*

$$f^t(\omega)(x) = \omega\big(f(x)\big). \tag{7.54}$$

pour tout $\omega \in F^*$ et $x \in E$.

Lemme 7.43.
Soit E muni de la base $\{e_i\}$ et F muni de la base $\{g_i\}$ et une application $f : E \to F$. Si A est la matrice de f dans ces bases, alors A^t est la matrice de f^t dans les bases $\{e_i^*\}$ et $\{g_i^*\}$ de E^* et F^*.

Démonstration. Nous allons montrer que les formes $f^t(g_i^*)$ et $\sum_k (A^t)_{ik} g_k^*$ sont égales en les appliquant à un vecteur.

Par définition de la matrice d'une application linéaire dans une base,

$$f^t(g_i^*) = \sum_j (f^t)_{ij} e^* j, \tag{7.55}$$

et

$$f(e_k) = \sum_l A_{kl} g_l. \tag{7.56}$$

Du coup, si $x = \sum_k x_k e_k$, nous avons

$$f^t(g_i^*)x = \sum_{kl} x_k g_i^* A_{kl} g_l = \sum_{kl} x_k A_{kl} \delta_{il} = \sum_k x_k A_{ki} = \sum_k (A^t)_{ik} x_k. \tag{7.57}$$

D'autre part,

$$\sum_k (A^t)_{ik} g_k^* x = \sum_{kl} (A^t)_{ik} g_k^* x_l e_l = \sum_k (A^t)_{ik} x_k. \tag{7.58}$$

Le fait que (7.57) et (7.58) donnent le même résultat prouve le lemme. □

En corollaire, les rangs de f et de f^t sont égaux parce que le rang est donné par la plus grande matrice carré de déterminant non nul. Nous prouvons cependant ce résultat de façon plus intrinsèque.

Lemme 7.44 (Gilles Dubois)**.**
Si $f : E \to F$ est une application linéaire, alors

$$\text{rang}(f) = \text{rang}(f^t). \tag{7.59}$$

Démonstration. Nous posons $\dim\ker(f) = p$ et donc $\operatorname{rang}(f) = n-p$. Soit $\{e_1, \dots, e_p\}$ une base de $\ker(f)$ que l'on complète en une base $\{e_1, \dots, e_n\}$ de E. Nous considérons maintenant les vecteurs

$$g_i = f(e_{p+i}) \tag{7.60}$$

pour $i = 1, \dots, n-p$. C'est à dire que les g_i sont les images des vecteurs qui ne sont pas dans le noyau de f. Prouvons qu'ils forment une famille libre. Si

$$\sum_{k=1}^{n-p} a_k f(e_{p+k}) = 0, \tag{7.61}$$

alors $f\big(\sum_k a_k e_{p+k}\big) = 0$, ce qui signifierait que $\sum_k a_k e_{p+k}$ se trouve dans le noyau de f, ce qui est impossible par construction de la base $\{e_i\}_{i=1,\dots,n}$. Étant donné que les vecteurs $g_1, \dots, g_{n-p}$ sont libres, nous les complétons en une base

$$\{\underbrace{g_1, \dots, g_{n-p}}_{\text{images}}, \underbrace{g_{n-p+1}, \dots, g_r}_{\text{complétion}}\} \tag{7.62}$$

de F.

Nous prouvons maintenant que $\operatorname{rang}(f^t) \geqslant n-p$ en montrant que les formes $\{g_i^*\}_{i=1,\dots,n-p}$ est libre (et donc l'espace image de f^f est au moins de dimension $n-p$). Pour cela nous prouvons que $f^t(g_i^*) = e_{i+p}^*$. En effet

$$f^t(g_i^*)e_k = g_i^*(fe_k), \tag{7.63}$$

Si $k = 1, \dots, p$, alors $fe_k = 0$ et donc $g_i^*(fe_k) = 0$; si $k = p + l$ alors

$$f^t(g_i^*)e_k = g_i^*(fe_{k+l}) = g_i^*(g_l) = \delta_{i,l} = \delta_{i,k-p} = \delta_{k,i+p}. \tag{7.64}$$

Donc $f^t(g_i^*) = e_{i+p}^*$. Cela prouve que les formes $f^t(g_i^*)$ sont libres et donc que

$$\operatorname{rang}(f^t) \geqslant n - p = \operatorname{rang}(f). \tag{7.65}$$

En appliquant le même raisonnement à f^t au lieu de f, nous trouvons

$$\operatorname{rang}\big((f^t)^t\big) \geqslant \operatorname{rang}(f^t) \tag{7.66}$$

et donc, vu que $(f^t)^t = f$, nous obtenons $\operatorname{rang}(f) = \operatorname{rang}(f^t)$.

□

Proposition 7.45 ([84])**.**
Si f est une application linéaire entre les espaces vectoriels E et F, alors nous avons

$$\operatorname{Image}(f^t) = \ker(f)^\perp. \tag{7.67}$$

Démonstration. Soient donc l'application $f\colon E \to F$ et sa transposée $f^t\colon F^* \to E^*$. Nous commençons par prouver que $\operatorname{Image}(f^t) \subset (\ker f)^\perp$. Pour cela nous prenons $\omega \in \operatorname{Image}(f^t)$, c'est à dire $\omega = \alpha \circ f$ pour un certain élément $\alpha \in F^*$. Si $z \in \ker(f)$, alors $\omega(z) = (\alpha \circ f)(z) = 0$, c'est à dire que $\omega \in (\ker f)^\perp$.

Pour prouver qu'il y a égalité, nous n'allons pas démontrer l'inclusion inverse, mais plutôt prouver que les dimensions sont égales. Après, on sait que si $A \subset B$ et si $\dim A = \dim B$, alors $A = B$. Nous avons

$$\begin{aligned}
\dim\big(\operatorname{Image}(f^t)\big) &= \operatorname{rang}(f^t) && && \text{(7.68a)}\\
&= \operatorname{rang}(f) && \text{lemme 7.44} && \text{(7.68b)}\\
&= \dim(E) - \dim\ker(f) && \text{théorème 7.23} && \text{(7.68c)}\\
&= \dim\big((\ker f)^\perp\big) && \text{proposition 7.41.} && \text{(7.68d)}
\end{aligned}$$

□

Lemme 7.46 ([83])**.**
Soit $\mathbb{K}$ un corps, E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie et une application linéaire $f\colon E \to F$. L'application f est injective si et seulement si sa transposée[4] *f^t est surjective.*

Démonstration. Supposons que f soit injective. Alors par le lemme 7.35, il existe $g\colon F \to E$ tel que $g \circ f = \mathrm{Id}\,|_E$. Nous avons alors aussi $(g \circ f)^t = \mathrm{Id}\,|_{E^*}$, mais $(g \circ f)^t = f^t \circ g^t$, donc f^t est surjective.

Inversement, nous supposons que $f^t\colon F^* \to E^*$ est surjective. Alors en nous souvenant que E et F sont de dimension finie et en faisons jouer les identifications $(f^t)^t = f$ et $(E^*)^* = E$ nous pouvons savons qu'il existe $s\colon E^* \to F^*$ tel que $f^t \circ s = \mathrm{Id}\,|_{E^*}$. En passant à la transposée,

$$s^t \circ f = \mathrm{Id}\,|_E, \tag{7.69}$$

qui implique que f est injective. □

7.4.3 Transposée : sans le dual

Problèmes et choses à faire

J'aimerais une confirmation de ce qui est écrit dans ici à propos du fait que si f est une application linéaire, f^t n'est pas bien définie.

Il est légitime, si $f\colon V \to V$ est une application linéaire, de dire que sa transposée soit l'application linéaire $f^t\colon V \to V$ dont la matrice est la matrice transposée de celle de f. Lorsque nous travaillons sur $\mathbb{R}^n$ muni de la base canonique, cela ne pose pas de problèmes et nous pouvons écrire des égalités du type $\langle x, Ay\rangle = \langle A^t x, y\rangle$.

Hélas nous allons voir que cette façon de définir une transposée est mauvaise. Nous reprenons les notations de 7.36. Nous avons

$$f(x) = \sum_{kl} A_{kl} x_l e_k = \sum_{\alpha\beta} A'_{\alpha\beta} x'_\beta e'_\alpha. \tag{7.70}$$

Nous écrivons l'application transposée par rapport à la première base :

$$f^{t_1}(x) = \sum_{kl} A_{lk} x_l e_k \tag{7.71}$$

et par rapport à la seconde base :

$$f^{t_2}(x) = \sum_{\alpha\beta} A'_{\beta\alpha} x'_\beta e'_\alpha. \tag{7.72}$$

En replaçant dans $f^{(t_1)}$ les nombres x_l et les vecteurs e_k par les valeurs en termes des x'_β et e'_α nous obtenons

$$f^{t_1}(x) = \sum_{\alpha\beta} (Q^{-1}A^tQ)_{\alpha\beta} x'_\beta e'_\alpha. \tag{7.73}$$

Nous avons égalité avec (7.72) seulement lorsque $A'_{\beta\alpha} = (Q^{-1}A^tQ)_{\alpha\beta}$, mais sachant que $A' = Q^{-1}AQ$ nous avons besoin que

$$Q^{-1} = Q^t \tag{7.74}$$

(cela étant une égalité de matrice).

7.47.
Autrement dit, la façon «usuelle» de voir la transposée d'une application linéaire ne fonctionne dans les livres pour enfant uniquement parce qu'on n'y considère toujours $\mathbb{R}^n$ muni de la base canonique ou de bases orthonormées.

Notons que nous avons tout de mêmes le notions d'opérateur adjoint et autoadjoint pour parler d'application orthogonale sans passer par la transposée, voir 9.47.

4. Définition 7.42.

7.4.4 Polynômes de Lagrange

Soit $E = \mathbb{R}_n[X]$ l'ensemble des polynômes à coefficients réels de degré au plus n. Soient les $n+1$ réels distincts $a_0, \dots, a_n$. Nous considérons les formes linéaires associées $f_i \in E^*$,

$$f_i(P) = P(a_i). \tag{7.75}$$

Lemme 7.48.
Ces formes forment une base de E^*.

Démonstration. Nous prouvons que l'orthogonal est réduit au nul :

$$\text{Span}\{f_0, \dots, f_n\}^\perp = \{0\} \tag{7.76}$$

pour que la proposition 7.41 conclue. Si $P \in \text{Span}\{f_i\}^\perp$, alors $f_i(P) = 0$ pour tout i, ce qui fait que $P(a_i) = 0$ pour tout $i = 0, \dots, n$. Un polynôme de degré au plus n qui s'annule en $n+1$ points est automatiquement le polynôme nul. □

Les **polynômes de Lagrange** sont les polynômes de la base (pré)duale de la base $\{f_i\}$.

Proposition 7.49.
Les polynômes de Lagrange sont donnés par

$$P_i = \prod_{k \neq i} \frac{X - a_k}{a_i - a_k}. \tag{7.77}$$

Démonstration. Il suffit de vérifier que $f_j(P_i) = \delta_{ij}$. Nous avons

$$f_j(P_i) = P_i(a_j) = \prod_{k \neq i} \frac{a_j - a_k}{a_i - a_k}. \tag{7.78}$$

Si $j \neq i$ alors un des termes est nul. Si au contraire $i = j$, tous les termes valent 1. □

7.4.5 Dual de $\mathbb{M}(n, \mathbb{K})$

Proposition 7.50 ([2])**.**
Soit $\mathbb{K}$, un corps. Les formes linéaires sur $\mathbb{M}(n, \mathbb{K})$ sont les applications de la forme

$$\begin{aligned} f_A \colon \mathbb{M}_n(\mathbb{K}) &\to \mathbb{K} \\ M &\mapsto \text{Tr}(AM). \end{aligned} \tag{7.79}$$

Démonstration. Nous considérons l'application

$$\begin{aligned} f \colon \mathbb{M}(n, \mathbb{K}) &\to \mathbb{M}(n, \mathbb{K})' \\ A &\mapsto f_A \end{aligned} \tag{7.80}$$

et nous voulons prouver que c'est une bijection. Étant donné que nous sommes en dimension finie, nous avons égalité des dimensions de $\mathbb{M}_n(\mathbb{K})$ et $\mathbb{M}_n(\mathbb{K})'$, et il suffit de prouver que f est injective. Soit donc A telle que $f_A = 0$. Nous l'appliquons à la matrice $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$:

$$0 = f_A(E_{ij}) = \sum_k (AE_{ij})_{kk} = \sum_{kl} A_{kl}\delta_{il}\delta_{jk} = A_{ij}. \tag{7.81}$$

Donc $A = 0$. □

Corollaire 7.51 ([2])**.**
Soit $\mathbb{K}$ un corps et $\phi \in \mathbb{M}(n, \mathbb{K})^*$ telle que pour tout $X, Y \in \mathbb{M}(n, \mathbb{K})$ on ait

$$\phi(XY) = \phi(YX). \tag{7.82}$$

Alors il existe $\lambda \in \mathbb{K}$ tel que $\phi = \lambda \,\text{Tr}$.

Démonstration. La proposition 7.50 nous donne une matrice $A \in \mathbb{M}(n,\mathbb{K})$ telle que $\phi = f_A$. L'hypothèse nous dit que $f_A(XY) = f_A(YX)$, c'est à dire

$$\mathrm{Tr}(AXY) = \mathrm{Tr}(AYX) \tag{7.83}$$

pour toute matrices $X, Y \in \mathbb{M}(n,\mathbb{K})$. L'invariance cyclique de la trace nous dit que $\mathrm{Tr}(AXY) = \mathrm{Tr}(XAY)$, ce qui signifie que

$$\mathrm{Tr}\big((AX - XA)Y\big) = 0 \tag{7.84}$$

ou encore que $f_{AX-XA} = 0$ pour tout X. La fonction f étant injective nous en déduisons que la matrice A doit satisfaire

$$AX = XA \tag{7.85}$$

pour tout $X \in \mathbb{M}$. En particulier en prenant la fameuse matrice E_{ij} et en calculant un peu,

$$A_{li}\delta_{jm} = \delta_{il}A_{jm} \tag{7.86}$$

pour tout i, j, l, m. Cela implique que $A_{ll} = A_{mm}$ pour tout l et m et que $A_{jm} = 0$ dès que $j \neq m$. Il existe donc $\lambda \in \mathbb{K}$ tel que $A = \lambda\mathbb{1}$. En fin de compte,

$$\phi(X) = f_{\lambda\mathbb{1}}(X) = \lambda\,\mathrm{Tr}(X). \tag{7.87}$$

□

Corollaire 7.52 ([2])**.**
Soit $\mathbb{K}$ *un corps. Tout hyperplan de* $\mathbb{M}(n,\mathbb{K})$ *coupe* $\mathrm{GL}(n,\mathbb{K})$*.*

Démonstration. Soit $\mathcal{H}$ un hyperplan de $\mathbb{M}$. Il existe une forme linéaire ϕ sur $\mathbb{M}(n,\mathbb{K})$ telle que $\mathcal{H} = \ker(\phi)$. Encore une fois la proposition 7.50 nous donne $A \in \mathbb{M}$ telle que $\phi = f_A$; nous notons r le rang de A. Par le lemme 7.28 nous avons $A = PJ_rQ$ avec $P, Q \in \mathrm{GL}(n,\mathbb{K})$ et

$$J_r = \begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}. \tag{7.88}$$

Pour tout $X \in \mathbb{M}$ nous avons

$$\phi(X) = \mathrm{Tr}(AX) = \mathrm{Tr}(PJ_rQX) = \mathrm{Tr}(J_rQXP). \tag{7.89}$$

Ce que nous cherchons est $X \in \mathrm{GL}(n,\mathbb{K})$ telle que $\phi(X) = 0$. Nous commençons par trouver $Y \in \mathrm{GL}(n,\mathbb{K})$ telle que $\mathrm{Tr}(J_rY) = 0$. Celle-là est facile : c'est

$$Y = \begin{pmatrix} 0 & 1 \\ \mathbb{1}_{n-1} & 0 \end{pmatrix}. \tag{7.90}$$

Les éléments diagonaux de J_rY sont tous nuls. Par conséquent en posant $X = Q^{-1}YP^{-1}$ nous avons notre matrice inversible dans le noyau de ϕ. □

Chapitre 8

Classification de certains groupes

8.1 Théorèmes de Sylow

Lemme 8.1.
Soient H et K des sous-groupes finis de G. Alors

$$\mathrm{Card}(HK) = \frac{|H| \cdot |K|}{|H \cap K|}. \tag{8.1}$$

Attention : dans ce lemme, l'ensemble HK n'est pas spécialement un groupe. Ce serait le cas si H normaliserait K, c'est à dire si nous avions $hkh^{-1} \in K <, \forall h, k \in H \times K$.

Théorème 8.2 (Théorème de Cauchy[85])**.**
Soit G un groupe fini et p un nombre premier divisant $|G|$. Alors

(1) G contient un élément d'ordre p.

(2) Si G est un p-groupe, il existe un élément central d'ordre p dans G.

Lemme 8.3 (Théorème de Cayley)**.**
Si G est un groupe d'ordre n alors il est isomorphe à un sous-groupe du groupe symétrique S_n.

Démonstration. L'action à gauche de G sur lui-même

$$\begin{aligned} \varphi\colon G &\to S_n \\ \varphi(x)g &\mapsto xg \end{aligned} \tag{8.2}$$

est une permutation des éléments de G. Cela donne un morphisme injectif parce que si $\varphi(x) = \varphi(y)$ nous avons $xg = yg$ pour tout g et en particulier pour $g = e$ nous trouvons $x = y$. □

Lemme 8.4.
Soit p un diviseur premier de n. Alors le groupe symétrique S_n se plonge dans $\mathrm{GL}_n(\mathbb{F}_p)$.

Démonstration. Soit $\{e_i\}$ la base canonique de $\mathbb{F}_p$. Nous avons le morphisme injectif $\varphi\colon S_n \to \mathrm{GL}(n, \mathbb{F})$ donné par $\varphi(\sigma)e_i = e_{\sigma(i)}$. □

Remarque 8.5.
En mettant bout à bout les lemmes 8.3 et 8.4, nous trouvons que si p est un diviseur premier de $|G|$, alors G peut être vu comme un sous-groupe de $\mathrm{GL}(n, \mathbb{F}_p)$.

Définition 8.6.
*Soit p un nombre premier. Un p-**groupe** est un groupe dont tous les éléments sont d'ordre p^m pour un certain m (dépendant de l'élément).*

*Soit G un groupe fini et p, un diviseur premier de $|G|$. Un p-**Sylow** dans G est un p-sous-groupe d'ordre p^n où p^n est la plus grande puissance de p divisant $|G|$.*

Notons que si p est un nombre premier, alors tout groupe d'ordre p^m est un p-groupe.

Lemme 8.7.
Soit G un groupe fini et P, Q des p-sous-groupes. Nous supposons que Q normalise P. Alors PQ est un p-sous-groupe de G.

Si S est un p-Sylow, alors p ne divise pas le nombre $|G : S| = |G|/|S|$.

Proposition 8.8.
Soit le corps fini $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ (p premier). Soit T le sous-ensemble de $\mathrm{GL}_n(\mathbb{F}_p)$ formé des matrices triangulaires supérieures de rang[1] *n et dont les éléments diagonaux sont 1. Alors T est un p-Sylow de $\mathrm{GL}_n(\mathbb{F}_p)$.*

Démonstration. Nous commençons par étudier le cardinal de $\mathrm{GL}_n(\mathbb{F}_p)$. Pour la première colonne, la seule contrainte à vérifier est qu'elle ne soit pas nulle. Il y a donc $p^n - 1$ possibilités. Pour la seconde, il faut ne pas être multiple de la première. Il y a donc $p^n - p$ possibilités (parce qu'il y a p multiples possibles de la premières colonne). Pour la k-ième colonne, il faut éviter toutes les combinaisons linéaires des $(k-1)$ premières colonnes. Il y a p^{k-1} telles combinaisons et donc $p^n - p^{k-1}$ possibilités pour la k-ième colonne. Nous avons donc

$$\begin{aligned}\mathrm{Card}\left(\mathrm{GL}(n,\mathbb{F}_p)\right) &= (p^n-1)(p^n-p)\dots(p^n-p^{n-1}) && \text{(8.3a)}\\ &= p\cdot p^2\cdots p^{n-1}(p^n-1)(p^{n-1}-1)\dots(p-1) && \text{(8.3b)}\\ &= p^{\frac{n(n-1)}{2}}m && \text{(8.3c)}\end{aligned}$$

où m est un entier qui ne divise pas p.

En ce qui concerne le cardinal de T, le calcul est plus simple : pour la première ligne nous avons p^{n-1} choix (parce qu'il y a un 1 qui est imposé sur la diagonale), pour la seconde p^{n-2}, etc. En tout nous avons alors

$$|T| = p^{\frac{n(n-1)}{2}}, \tag{8.4}$$

et T est un p-Sylow de $\mathrm{GL}_n(\mathbb{F}_p)$. □

Proposition 8.9.
Soit p un nombre premier. Un groupe fini G est un p-groupe si et seulement l'ordre de G est p^n pour un certain n.

Démonstration. Supposons que G est un p-groupe. Soit q un nombre premier divisant $|G|$. Par le théorème de Cauchy (8.2), le groupe G contient un élément d'ordre q, soit g un tel élément. Étant donné que G est un p-groupe, $g^{p^n} = g^q = e$ pour un certain n. Donc $q = p^n$ et $q = p$ parce que q est premier. Nous venons de prouver que p est le seul nombre premier qui divise $|G|$. L'ordre de G est par conséquent une puissance de p.

Nous nous intéressons maintenant à l'implication inverse. Nous supposons que $|G| = p^n$ pour un certain entier $n \geqslant 0$. Soit $g \in G$; nous notons r l'ordre de G. Le sous-groupe $\mathrm{gr}(g)$ est d'ordre r, donc r divise $|G|$ (par le théorème 3.54 de Lagrange). Le nombre r est alors une puissance de p. □

Lemme 8.10.
Soit G, un groupe fini de cardinal $|G| = n$ et p, un diviseur premier de n. Nous notons $n = p^m \cdot r$ où p ne divise pas r. Soit H un sous-groupe de G et S, un p-Sylow de G. Alors il existe $g \in G$ tel que

$$gSg^{-1} \cap H \tag{8.5}$$

soit un p-Sylow de H.

1. Définition 7.22.

Démonstration. Nous considérons l'ensemble G/S sur lequel H agit. Si $a \in G$, le stabilisateur de $[a]$ dans G/S est

$$\begin{aligned} \text{Fix}\big([a]\big) &= \{h \in H \text{ tel que } [ha] = [a]\} && (8.6a)\\ &= \{h \in H \text{ tel que } a^{-1}ha \in S\} && (8.6b)\\ &= aSa^{-1} \cap H. && (8.6c)\end{aligned}$$

Nous cherchons $a \in G$ tel que l'entier

$$\frac{\text{Card}(H)}{\text{Card}\big(aSa^{-1} \cap H\big)} \tag{8.7}$$

soit premier avec p. En effet, dans ce cas le groupe $\text{Fix}([a])$ est un p-Sylow de H parce que $|H : aSa^{-1} \cap H|$ ne divise pas p. La formule des orbites (équation (3.139)) nous dit que

$$\frac{|H|}{|aSa^{-1} \cap H|} = \text{Card}\big(\mathcal{O}_{[a]}\big). \tag{8.8}$$

Supposons que toutes les orbites aient un cardinal divisible par p. Étant donné que G/S est une réunion disjointe de ses orbites, nous aurions

$$p \mid \text{Card}(G/S) = \frac{|G|}{|S|} \tag{8.9}$$

alors que S étant un p-Sylow, p ne peut pas diviser $|G|/|S|$. Toutes les orbites n'ont donc pas un cardinal divisible par p, et il existe un $a \in G$ tel que (8.7) soit vérifiée. □

Théorème 8.11 (Théorème de Sylow)**.**
Soit G un groupe fini et p, un diviseur premier de $|G|$. Alors

(1) G possède au moins un p-Sylow[2]*.*

(2) Tout p-sous-groupe de G est contenu dans un p-Sylow.

(3) Les p-Sylow de G sont conjugués.

(4) Si n_p est le nombre de p-Sylow de G, alors n_p divise $|G|$ et $n_p \in [1]_p$.

Démonstration.

(1) Nous savons de la remarque 8.5 que G est un sous-groupe de $\text{GL}_n(\mathbb{F}_p)$ et que ce dernier a un p-Sylow par la proposition 8.8. Par conséquent G possède un p-Sylow par le lemme 8.10.

(2) Soit H un p-sous-groupe de G et S, un p-Sylow de G (qui existe par le point précédent). Par le lemme 8.10 il existe $a \in G$ tel que $aSa^{-1} \cap H$ soit un p-Sylow de H. Mais H est un p-groupe et un p-Sylow dans un p-groupe est automatiquement le groupe entier. Par conséquent,

$$H = aSa^{-1} \cap H \tag{8.10}$$

et $H \subset aSa^{-1}$, ce qui signifie que H est inclus à un p-Sylow.

(3) Soit H un p-Sylow. Nous venons de voir que si S est un p-Sylow quelconque, alors H est inclus au p-Sylow aSa^{-1} pour un certain $a \in G$. Donc H est un p-Sylow inclus dans le p-Sylow aSa^{-1}, donc $H = aSa^{-1}$.

(4) Le fait que n_p divise n est parce que tous les p-Sylow ont le même nombre d'éléments (ils sont conjugués) et sont deux à deux disjoints. Donc ils forment une partition de G et $|G| = n_p|S|$ si S est un p-Sylow quelconque.

Montrons maintenant que n_p est congru à un modulo p. Soit E l'ensemble des p-Sylow de G. Le groupe G agit sur E par conjugaison. Soit S un p-Sylow et considérons l'ensemble

$$E_S = \{T \in E \text{ tel que } s \cdot T = T \forall s \in S\}. \tag{8.11}$$

2. Définition 8.6.

où l'action est celle par conjugaison. C'est l'ensemble des points fixes de E sous l'action de S. L'ensemble E est la réunion des orbites sous S et chacune de ces orbites a un cardinal qui divise $|S| = p^m$. Par conséquent $|\mathcal{O}_T|$ vaut 1 lorsque $T \in E_S$ et est un multiple de p sinon. Nous avons donc

$$|E| \equiv |E_S| \mod p. \tag{8.12}$$

Nous voulons obtenir $|E_S| = 1$. Évidemment $S \in E_S$ parce que si $s \in S$ alors $sSs^{-1} = S$. Nous voudrions montrer que S est le seul élément de E_S. Soit $T \in E_S$, c'est à dire que T est un p-Sylow de G tel que

$$sTs^{-1} = T \tag{8.13}$$

pour tout $s \in S$. Soit N le groupe engendré par S et T. Montrons que T est normal dans N. Un élément g dans N s'écrit

$$g = s_1 t_1 \cdots s_r t_r \tag{8.14}$$

avec $s_i \in S$ et $t_i \in T$. Si $t \in T$, en utilisant le fait que T est un groupe et le fait que S le normalise, nous avons

$$gtg^{-1} = s_1 t_1 \dots s_r t_r t t_r^{-1} s_r^{-1} \dots t_1^{-1} s_r^{-1} \in T. \tag{8.15}$$

Donc T est un sous-groupe normal de N. Mais S et T sont conjugués dans N (parce que ils sont des p-Sylow de N), donc il existe un élément $a \in N$ tel que $aTa^{-1} = S$. Mais étant donné que T est normal,

$$S = aTa^{-1} = T. \tag{8.16}$$

Ceci achève la démonstration des théorèmes de Sylow.

□

Proposition 8.12.
Si S est un p-Sylow dans le groupe G alors pour tout $g \in G$, l'ensemble gSg^{-1} est encore un p-groupe.

Démonstration. Si les éléments de S sont d'ordre p^n, alors nous avons

$$(gsg^{-1})^q = gs^q g^{-1} = e. \tag{8.17}$$

Pour avoir $gs^q g^{-1} = e$, il faut et suffit que $gs^q = g$, alors $s^q = e$, c'est à dire $q = p^n$. Donc gSg^{-1} est encore un p-Sylow. □

Lemme 8.13 ([86])**.**
Soit G, un groupe fini et p, un nombre premier. Si H et K sont des groupes distincts d'ordre p, alors $H \cap K = \{e\}$.

Démonstration. L'ensemble $H \cap K$ est un sous-groupe de H. Par conséquent son ordre divise celui de H qui est un nombre premier. Par conséquent soit $|H \cap K| = 1$, soit $|H \cap K| = |H|$. Dans le second cas nous aurions $H = K$, alors que nous avons supposé que H et K étaient distincts. □

Proposition 8.14 ([86])**.**
Soit G un groupe fini et n le nombre de sous-groupes d'ordre p dans G. Alors le nombre d'éléments d'ordre p dans G vaut $n(p-1)$.

Démonstration. Si g est un élément d'ordre p dans G, le groupe H engendré par g est d'ordre p. Réciproquement si H est un groupe d'ordre p, tous les éléments de $H\backslash\{e\}$ sont d'ordre p (parce que l'ordre d'un élément divise l'ordre du groupe). Donc l'ensemble des éléments d'ordre p dans G est la réunion des ensembles $H\backslash\{e\}$ où H parcours les sous-groupes d'ordre p dans G. Chacun de ces ensembles possède $p-1$ éléments et le lemme 8.13 nous assure qu'ils sont disjoints. Par conséquent nous avons $n(p-1)$ éléments d'ordre p dans G. □

Corollaire 8.15.
Un groupe d'ordre premier est cyclique.

Démonstration. Soit p l'ordre de G. Le nombre de sous-groupes d'ordre p est $n = 1$ (et c'est G lui-même). La proposition 8.14 nous dit alors que le nombre d'éléments d'ordre p dans G est $p-1$. Donc tout élément est générateur. □

Lemme 8.16.
Le groupe A_6 n'accepte pas de sous-groupes normaux d'ordre 60.

Démonstration. Soit G normal dans A_6, et a, un élément d'ordre 5 dans G (qui existe parce que 5 divise 60). Soit aussi un élément b d'ordre 5 dans A_6. Les groupes $\text{gr}(a)$ et $\text{gr}(b)$ sont deux 5-Sylow dans A_6. En effet, 5 un nombre premier et est la plus grande puissance de 5 dans la décomposition de 60 ; donc $\text{gr}(a)$ est un 5-Sylow dans G. D'autre part, l'ordre de A_6 (qui est $\frac{1}{2}6!$) ne possède également que 5 à la puissance 1 dans sa décomposition.

En vertu du théorème de Sylow 8.11(3), les 5-Sylow $\text{gr}(a)$ et $\text{gr}(b)$ sont conjugués et il existe $\tau \in A_6$ tel que $b = \tau a \tau^{-1}$. Mais G étant normal dans A_6, l'élément $\tau a \tau^{-1}$ est encore dans G, de telle sorte que $b \in G$. Du coup G doit contenir tous les éléments d'ordre 5 de A_6.

Les éléments d'ordre 5 de A_6 doivent fixer un des points de $\{1, 2, 3, 4, 5, 6\}$ puis permuter les autres de façon à n'avoir qu'un seul cycle. Un cycle correspond à écrire les nombres $1, 2, 3, 4, 5$ dans un certain ordre. Ce faisant, le premier n'a pas d'importance parce qu'on considère la permutation cyclique, par exemple $(3, 5, 2, 1, 4)$ est la même chose que $(5, 2, 1, 4, 3)$. Le nombre de cycles sur $\{1, 2, 3, 4, 5\}$ est donc de 4!, et par conséquent le nombre d'éléments d'ordre 5 dans A_6 est $6 \cdot 4! = 144$.

Le groupe G doit contenir au moins 144 éléments alors que par hypothèse il en contient 60 ; contradiction.

□

8.2 Un peu de classification de groupes

8.2.1 Groupe monogène

Théorème 8.17.
Un groupe monogène est abélien. Plus précisément,

(1) un groupe monogène infini est isomorphe à $\mathbb{Z}$,

(2) un groupe monogène fini est isomorphe à $\mathbb{Z}/n\mathbb{Z}$ pour un certain n.

Un groupe monogène d'ordre n possède $\varphi(n)$ générateur où φ est la fonction indicatrice d'Euler définie en 3.104.

Démonstration. Le groupe est abélien parce que $g = a^n$, $g' = a^{n'}$ implique $gg' = q^{n+n'} = g'g$. Nous considérons un générateur a de G (qui existe parce que G est monogène) et le morphisme surjectif

$$\begin{aligned} f\colon \mathbb{Z} &\to G \\ p &\mapsto a^p. \end{aligned} \tag{8.18}$$

Si G est infini, alors f est injective parce que si $a^n = a^{n'}$, alors $a^{n-n'} = e$, ce qui rendrait G cyclique et par conséquent non infini. Nous concluons que si G est infini, alors f est une bijection et donc un isomorphisme $\mathbb{Z} \simeq G$.

Si G est fini, alors f n'est pas injective et a un noyau $\ker f$. Étant donné que $\ker f$ est un sous-groupe de G, il existe un (unique) n tel que $\ker f = n\mathbb{Z}$ et le premier théorème d'isomorphisme (théorème 3.22) nous indique que

$$\mathbb{Z}/\ker f = \mathbb{Z}/n\mathbb{Z} = \text{Image}\, f = G. \tag{8.19}$$

Dans ce cas, le fait qu'un groupe monogène d'ordre n possède $\varphi(n)$ générateurs est le contenu de la proposition 3.107. □

Le lemme suivant donne une démonstration alternative, avec une construction plus explicite de l'isomorphisme.

Lemme 8.18 ([1])**.**
À props de groupes monogènes[3]

(1) *Soit un groupe monogène G d'ordre fini n dont g est un générateur. Alors il existe un isomorphisme*
$$\phi\colon G \to (\mathbb{Z}/n\mathbb{Z}, +) \tag{8.20}$$
tel que $\phi(g) = 1$.

(2) *Si G est un groupe monogène d'ordre infini et si g est un générateur, alors il existe un isomorphisme*
$$\phi\colon G \to (\mathbb{Z}, +) \tag{8.21}$$
tel que $\phi(g) = 1$.

(3) *Soient G et H deux groupes monogènes de même ordre. Soient g un générateur de G et h, un générateur de H. Il existe un isomorphisme de G sur H qui envoie g sur h.*

Démonstration. Commençons par enfoncer une porte ouverte : vu que le groupe est monogène, l'ordre du groupe est égal à l'ordre de son générateur. Nous séparons les cas selon quel l'ordre soit fini ou non.

L'ordre de G est fini et vaut n Si $k \in \mathbb{Z}$, nous notons $[k]_n$ la classe de k modulo n, c'est à dire l'ensemble $\{k + pn \text{ tel que } p \in \mathbb{Z}\}$.
Nous construisons l'isomorphisme $\phi\colon G \to \mathbb{Z}/n\mathbb{Z}$ de la façon suivante :
$$\phi(g^m) = [m]_n. \tag{8.22}$$
Cela est une bonne définition parce que une égalité du type $g^m = g^{m'}$ implique que m et m' soient dans la même classe modulo n. Nous vérifions que cela est une isomorphisme entre G et $\mathbb{Z}/n\mathbb{Z}$.

Morphisme Pour l'identité, si $x = e$ alors $m = 0$ et $\phi(e) = [0]_n$. Et si $x = g^k$, $y = g^l$ alors $\phi(xy) = \phi(g^{k+l}) = [k + l]_n = [k]_n + [l]_n = \phi(x) + \phi(y)$.

Injectif Supposons $\phi(g^k) = \phi(g^l)$ avec $k \geqslant l$. Nous avons $h^k = h^l$, dont $h^{k-l} = e$, ce qui donne $k - l \in [0]_n$ ou encore $[k]_n = [l]_n$. En particulier $g^k = g^l$.

Surjectif La classe $[k]_n$ est l'image de g^k.

L'ordre de G est infini Si l'ordre de G est infini alors un élément $x \in G$ s'écrit de façon unique sous la forme $x = g^m$ avec $m \in \mathbb{Z}$. Dans ce cas nous définissons directement $\phi(g^m) = m$.
Le reste de la preuve est alors identique au cas d'ordre fini, mais sans les complications liées au modulo.

La dernière assertion s'obtient des précédentes par composition d'isomorphismes.

□

8.2.2 Automorphismes du groupe $\mathbb{Z}/n\mathbb{Z}$

Notons que $\mathbb{Z}/n\mathbb{Z} = \mathbb{F}_n$ est un groupe pour l'addition tandis que $(\mathbb{Z}/n\mathbb{Z})^*$ est un groupe pour la multiplication. Il ne peut donc pas y avoir d'équivoque.

Théorème 8.19 ([87])**.**
Pour chaque $x \in (\mathbb{Z}/n\mathbb{Z})^*$ nous considérons l'application
$$\begin{aligned} \sigma_x\colon \mathbb{Z}/n\mathbb{Z} &\to \mathbb{Z}/n\mathbb{Z} \\ y &\mapsto xy. \end{aligned} \tag{8.23}$$
L'application
$$\sigma\colon \big((\mathbb{Z}/n\mathbb{Z})^*,\ \cdot\ \big) \to \operatorname{Aut}\big(\mathbb{Z}/n\mathbb{Z}, +\big) \tag{8.24}$$
ainsi définie est un isomorphisme de groupes.

3. Définition 3.6.

L'énoncé de ce théorème s'écrit souvent rapidement par

$$\mathrm{Aut}(\mathbb{Z}/n\mathbb{Z}) = (\mathbb{Z}/n\mathbb{Z})^*, \tag{8.25}$$

mais il faut bien garder à l'esprit qu'à gauche on considère le groupe additif et à droite celui multiplicatif.

Démonstration. Nous notons $[x]$ la classe de x dans $\mathbb{Z}/n\mathbb{Z}$. Nous avons $\mathbb{Z}/n\mathbb{Z} = [1]$. Soit f un automorphisme de $(\mathbb{Z}/n\mathbb{Z}, +)$; pour tout $r \in \mathbb{Z}$ nous avons

$$f([r]) = f(r[1]) = rf([1]) = [r]f([1]). \tag{8.26}$$

En particulier, vu que f est surjective, il existe un r tel que $f([r]) = [1]$. Pour un tel r nous avons $[1] = [r]f([1])$, c'est à dire que nous avons montré que $f([1])$ est inversible dans $\big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\big)$. Nous montrons à présent que [4]

$$\begin{aligned} \sigma\colon \mathrm{Aut}((\mathbb{Z}/n\mathbb{Z}, +)) &\to \big((\mathbb{Z}/n\mathbb{Z})^*, \cdot\big) \\ f &\mapsto f([1]) \end{aligned} \tag{8.27}$$

est un isomorphisme.

Nous commençons par la surjectivité. Soit $[a] \in (\mathbb{Z}/n\mathbb{Z})^*$. Les élément $[a]$ et $[1]$ étant tous deux des générateurs de $(\mathbb{Z}/n\mathbb{Z}, +)$, il existe un automorphisme de $\mathbb{Z}/n\mathbb{Z}$ qui envoie $[1]$ sur $[a]$ par le lemme 8.18. Cela prouve la surjectivité de σ.

En ce qui concerne l'injectivité, considérons des automorphismes f_1 et f_2 de $(\mathbb{Z}/n\mathbb{Z}, +)$ tels que $f_1([1]) = f_2([1])$. Les automorphismes f_1 et f_2 prennent la même valeur sur un générateur et donc sur tout le groupe. Donc $f_1 = f_2$.

Enfin nous prouvons que σ est un morphisme, c'est à dire que $\sigma(f \circ g) = \sigma(f)\sigma(g)$. Nous avons

$$f\big(g([1])\big) = f\big(g([1])[1j]\big) = g([1])f([1]) = \sigma(f)\sigma(g). \tag{8.28a}$$

□

Ce dernier résultat s'étend aux groupes cycliques.

Proposition 8.20.
Si G est un groupe cyclique d'ordre n, alors

$$\mathrm{Aut}(G) = (\mathbb{Z}/n\mathbb{Z})^*. \tag{8.29}$$

Corollaire 8.21.
Si p divise $q-1$ alors $\mathrm{Aut}(\mathbb{F}_q)$ *possède un unique sous-groupe d'ordre p.*

Démonstration. Si a est un générateur de $\mathbb{F}_q^*$ alors le groupe

$$\mathrm{gr}\left(a^{\frac{q-1}{p}}\right) \tag{8.30}$$

est un sous-groupe d'ordre p. En ce qui concerne l'unicité, soit S un sous-groupe d'ordre p. Il est donc d'indice $(q-1)/p$ dans $\mathbb{F}_q^*$ et le lemme 3.60 nous enseigne que le groupe donné en (8.30) est contenu dans S. Il est donc égal à S parce qu'il a l'ordre de S. Le fait que S soit normal est dû au fait que $\mathbb{F}_q^*$ est abélien. □

4. Le σ donné ici est l'inverse de celui donné dans l'énoncé. Cela ne change évidemment rien à la validité de l'énoncé et de la preuve.

8.2.3 Groupes abéliens finis

Source : [25].

Nous rappelons que l'exposant d'un groupe fini est le ppcm des ordres de ses éléments. Dans le cas des groupes abéliens finis, l'exposant joue un rôle important du fait qu'il existe un élément dont l'ordre est l'exposant. Cela est le théorème suivant.

Théorème 8.22 (Exposant dans un groupe abélien fini)**.**
Un groupe abélien fini contient un élément dont l'ordre est l'exposant du groupe.

Démonstration. Soit G un groupe abélien fini et $x \in G$, un élément d'ordre maximum m. Nous montrons par l'absurde que l'ordre de tous les éléments de G divise m. Soit donc $y \in G$, un élément dont l'ordre ne divise pas m ; nous notons q son ordre. Vu que q ne divise pas m, le nombre q possède au moins un facteur premier plus de fois que m : soit p premier tel que la décomposition de q contienne p^β et celle de m contienne p^α avec $\beta > \alpha$. Autrement dit,

$$m = p^\alpha m' \tag{8.31a}$$
$$q = p^\beta q' \tag{8.31b}$$

où m' et q' ne contiennent plus le facteur p. L'élément x étant d'ordre m, l'élément x^{p^α} est d'ordre m'. De la même manière, l'élément $y^{q'}$ est d'ordre p^β. Étant donné que p^β et m' sont premiers entre eux, l'élément $x^{p^\alpha}y^{q'}$ est d'ordre $p^\alpha m' > m$. D'où une contradiction avec le fait que x était d'ordre maximal.

Par conséquent l'ordre de tous les éléments de G divise celui de x qui est alors le ppcm des ordres de tous les éléments de G, c'est à dire l'exposant de G. □

Proposition 8.23.
Soit G un groupe abélien fini et $x \in G$, un élément d'ordre maximum. Alors

(1) Il existe un morphisme $\varphi\colon G \to \mathrm{gr}(x)$ tel que $\varphi(x) = x$.

(2) Il existe un sous-groupe K de G tel que $G = \mathrm{gr}(x) \oplus K$.

Démonstration. Nous notons a l'ordre de x qui est également l'exposant du groupe G.

Nous allons prouver la première partie par récurrence sur l'ordre du groupe. Si $G = \mathrm{gr}(x)$, alors c'est évident. Soit H un sous-groupe propre de G contenant x et tel que le problème soit déjà résolu pour H : il existe un morphisme $\varphi\colon H \to \mathrm{gr}(x)$ tel que $\varphi(x) = x$. Soit $y \in G\backslash H$, d'ordre b. Nous allons trouver un morphisme $\hat{\varphi}\colon \mathrm{gr}(H, y) \to \mathrm{gr}(x)$ telle que $\hat{\varphi}(x) = x$.

Pour cela nous commençons par construire les applications suivantes :

$$\begin{aligned} \tilde{\varphi}\colon \mathbb{Z}/b\mathbb{Z} \times H &\to \mathrm{gr}(x) \\ (\bar{k}, h) &\mapsto x^{kl}\varphi(h) \end{aligned} \tag{8.32}$$

où l est encore à déterminer, et

$$\begin{aligned} p\colon \mathbb{Z}/b\mathbb{Z} \times H &\to \mathrm{gr}(y, H) \\ (\bar{k}, h) &\mapsto y^k h. \end{aligned} \tag{8.33}$$

Pour que $\tilde{\varphi}$ soit bien définie, il faut que a divise bl. L'application p est bien définie parce que $\bar{k}$ est pris dans $\mathbb{Z}/b\mathbb{Z}$ et que b est l'ordre de y.

Nous allons construire le morphisme $\hat{\varphi}$ en considérant le diagramme

$$\begin{array}{ccccc} \ker(p) & \hookrightarrow & \mathbb{Z}/b\mathbb{Z} \times H & \xrightarrow{p} & \mathrm{gr}(y, H) \\ & & \downarrow{\scriptstyle \tilde{\varphi}} & \swarrow{\scriptstyle \hat{\varphi}} & \\ & & \mathrm{gr}(x) & & \end{array} \tag{8.34}$$

que l'on voudra être commutatif. Vu que p est surjective, les théorèmes d'isomorphismes nous disent que

$$\mathrm{gr}(y, H) \simeq \frac{\mathbb{Z}/b\mathbb{Z} \times H}{\ker p}. \tag{8.35}$$

Si $[\bar{k},h]$ est la classe de $(\bar{k},h)$ modulo $\ker(p)$ alors nous voudrions définir $\hat{\varphi}$ par

$$\hat{\varphi}\big([\bar{k},h]\big)=\tilde{\varphi}(\bar{k},h). \tag{8.36}$$

Pour que cela soit bien définit, il faut que si $(\bar{r},z)\in\ker p$,

$$\hat{\varphi}\big([\bar{k}\bar{r},hz]\big)=\hat{\varphi}\big([\bar{k},h]\big), \tag{8.37}$$

c'est à dire que $\tilde{\varphi}(\bar{r},z)=e$. Du coup la définition (8.36) n'est bonne que si et seulement si

$$\ker(p)\subset\ker(\tilde{\varphi}). \tag{8.38}$$

Nous pouvons obtenir cela en choisissant bien l.

Déterminons d'abord le noyau de p. Pour cela nous considérons un nombre β divisant b tel que $\operatorname{gr}(y)\cap H=\operatorname{gr}(y^\beta)$. Nous aurons $p(\bar{k},h)=e$ si et seulement si $y^h=e$. En particulier $h=y^{-k}\in\operatorname{gr}(y)\cap H=\operatorname{gr}(y^\beta)$. Si $h=(y^\beta)^m=y^{m\beta}$, alors $k=-m\beta$ et nous avons

$$\ker(p)=\{(-m\beta,y^{m\beta})\text{ tel que }m\in\mathbb{Z}\}. \tag{8.39}$$

En plus court : $\ker(p)=\operatorname{gr}(\beta,y^{-\beta})$. Nous devons donc fixer l de telle sorte que $\tilde{\varphi}(\beta,y^{-\beta})=e$. Étant donné que φ prend ses valeurs dans $\operatorname{gr}(x)$, il existe un entier α tel que $\varphi(y^{-\beta})=x^\alpha$; en utilisant cet α, nous écrivons

$$\tilde{\varphi}(\beta,y^{-\beta})=x^{\beta l}\varphi(y^{-\beta})=x^{\beta l+\alpha}. \tag{8.40}$$

Par conséquent nous choisissons $l=-\alpha/\beta$. Nous devons maintenant vérifier que ce choix est légitime, c'est à dire que a divise bl et que α/β est un entier.

Étant donné que y est d'ordre b,

$$e=\varphi(y^b)=\varphi(y^{-\beta b/\beta})=\varphi(y^{-\beta})^{b/\beta}=x^{b\beta/\alpha}. \tag{8.41}$$

Par conséquent a divise $\frac{b\alpha}{\beta}=-bl$.

Pour voir que l est entier, nous nous rappelons que a est l'exposant de G (parce que x est d'ordre maximum) et que par conséquent b divise a. Mais a divise $\alpha\frac{b}{\beta}$. Donc α/β est entier.

Nous passons maintenant à la seconde partie de la preuve. Nous considérons un morphisme $\varphi\colon G\to\operatorname{gr}(x)$ tel que $\varphi(x)=x$. La première partie nous en assure l'existence. Nous montrons que

$$\begin{aligned}\psi\colon G&\to\operatorname{gr}(x)\oplus\ker(\varphi)\\ g&\mapsto\big(\varphi(g),g\varphi(g)^{-1}\big)\end{aligned} \tag{8.42}$$

est un isomorphisme. D'abord $g\varphi(g)^{-1}$ est dans le noyau de φ parce que $\varphi(g)^{-1}$ étant dans $\operatorname{gr}(x)$, et φ étant un morphisme,

$$\varphi\big(g\varphi(g)^{-1}\big)=\varphi(g)\varphi(g)^{-1}=e. \tag{8.43}$$

L'application ψ est un morphisme parce que, en utilisant le fait que G est abélien,

$$\begin{aligned}\psi(g_1g_2)&=\big(\varphi(g_1g_2),g_1g_2\varphi(g_1g_2)^{-1}\big) &\text{(8.44a)}\\ &=\big(\varphi(g_1)\varphi(g_2),g_1\varphi(g_1)^{-1}g_2\varphi(g_2)^{-1}\big) &\text{(8.44b)}\\ &=\psi(g_1)\psi(g_2). &\text{(8.44c)}\end{aligned}$$

L'application ψ est injective parce que si $\psi(g)=(e,e)$ alors $\varphi(g)=e$ et $g\varphi(g)^{-1}=e$, ce qui implique $g=e$.

Enfin ψ est surjective parce qu'elle est injective et que les ensembles de départ et d'arrivée ont même cardinal. En effet par le premier théorème d'isomorphisme (théorème 3.22) appliqué à φ nous avons

$$|G|=|\operatorname{gr}(x)|\cdot|\ker(\varphi)|. \tag{8.45}$$

□

Théorème 8.24.
Tout groupe abélien fini (non trivial) se décompose en

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \oplus \ldots \oplus \mathbb{Z}/d_r\mathbb{Z} \tag{8.46}$$

avec $d_1 \geqslant 1$ et d_i divise d_{i+1} pour tout $i = 1, \ldots, r-1$.

De plus la liste $(d_1, \ldots, d_r)$ vérifiant ces propriétés est unique.

Démonstration. Soit x_1 un élément d'ordre maximal dans G. Soit n_1 son ordre et

$$H_1 = \mathrm{gr}(x_1) = \mathbb{F}_{n_1}. \tag{8.47}$$

D'après la proposition 8.23(2), il existe un supplémentaire K_1 tel que $G = \mathbb{F}_{n_1} \oplus K_1$. Si $K_1 = \{1\}$ on s'arrête et on garde $G = \mathbb{F}_{n_1}$. Sinon on continue de la sorte en prenant x_2 d'ordre maximal dans K_1 etc.

Nous devons maintenant prouver l'unicité de cette décomposition. Soit

$$G = \mathbb{F}_{d_1} \oplus \ldots \oplus \mathbb{F}_{d_r} = \mathbb{F}_{s_1} \oplus \ldots \oplus \mathbb{F}_{s_q}. \tag{8.48}$$

L'exposant de G est d_r et s_q. Donc $d_r = s_q$. Les complémentaires étant égaux nous avons

$$\mathbb{F}_{d_1} \oplus \ldots \oplus \mathbb{F}_{d_{r-1}} = \mathbb{F}_{s_1} \oplus \ldots \oplus \mathbb{F}_{s_{q-1}}. \tag{8.49}$$

En continuant nous trouvons $r = q$ et $d_i = s_i$. □

8.2.4 Groupes d'ordre *pq*

Lemme 8.25.
Soit G un groupe d'ordre pq où p et q sont des nombres premiers distincts. Nous supposons que $p < q$.

(1) Le groupe G possède un unique q-Sylow.

(2) Cet unique q-Sylow est normal dans G.

(3) Il n'est ni $\{e\}$ ni G.

(4) Le groupe G n'est pas un groupe simple[5].

Démonstration. Soit n_q le nombre de q-Sylow ; par le théorème de Sylow 8.11(1) le groupe G possède des q-Sylow et par 8.11(4),

$$n_q \in [1]_q. \tag{8.50}$$

De plus le nombre n_q divise $|G| = pq$. Donc n_q vaut p, q ou 1. Avoir $n_q = p$ n'est pas possible parce que $n_q \in [1]_q$ et $p < q$. Avoir $n_q = q$ n'est pas possible non plus, pour la même raison. Donc $n_q = 1$. Notons H l'unique q-Sylow de G.

Le fait que H soit normal est une conséquence de 8.11(3) parce que le conjugué de H est encore un q-Sylow alors que H est l'unique q-Sylow.

Vu que

$$1 < p = |H| < pq = |G|, \tag{8.51}$$

le sous-groupe H n'est ni réduit à l'identité ni le groupe entier.

Par conséquent G n'est pas simple parce qu'il contient un sous-groupe normal non trivial. □

Avant le lire le théorème suivant, n'oubliez pas de lire la définition d'un produit semi-direct 3.93.

Théorème 8.26 ([88])**.**
Soient deux nombres premiers distincts[6] *p et q avec $q > p$.*

5. Pas de sous-groupe normaux non triviaux, 3.5.
6. Le cas $p = q$ sera traité par la proposition 8.29.

(1) *Si p ne divise pas $q-1$ alors tout groupe d'ordre pq est cyclique et plus précisément le seul groupe (à isomorphisme près) d'ordre pq est $\mathbb{Z}/pq\mathbb{Z}$.*

(2) *Si $p \mid q-1$, alors il n'existe que deux groupes d'ordre pq :*
   — *Le groupe abélien et cyclique $\mathbb{Z}/pq\mathbb{Z}$.*
   — *Le produit semi-direct non abélien*

$$G = \mathbb{Z}/q\mathbb{Z} \times_\varphi \mathbb{Z}/p\mathbb{Z} \tag{8.52}$$

   où $\varphi(\bar{1})$ est d'ordre p dans $\mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$.

(3) *Si p et q sont premiers entre eux, le produit est direct*[7] .

Démonstration. Division de la preuve en plusieurs parties.

Préliminaires avec Sylow Soit un groupe G d'ordre pq. Soient H, un q-Sylow et K, un p-Sylow de G. Ils existent parce que p et q sont des diviseurs premiers de $|G|$ (théorème de Sylow 8.11). Si n_q est le nombre de q-Sylow dans G alors n_q divise $|G|$ et $n_q = 1 \mod q$. Donc d'abord n_q vaut 1, p ou q. Ensuite $n_q = q$ est exclu par la condition $n_q = 1 \mod q$; la possibilité $n_q = p$ est également impossible parce que $p = 1 \mod q$ est impossible avec $p < q$. Donc $n_q = 1$ et H est normal dans G.

L'ensemble $H \cap K$ est un sous-groupe à la fois de H et de K, ce qui entraine que (théorème de Lagrange 3.54) $|H \cap K|$ divise à la fois p et q. Nous en déduisons que $|H \cap K| = 1$ et donc que $H \cap K = \{e\}$.

Étant donné que H est normal, l'ensemble HK est un sous-groupe de G. De plus l'application

$$\begin{aligned} \psi\colon H \times K &\to HK \\ (h,k) &\mapsto hk \end{aligned} \tag{8.53}$$

est un bijection. Nous ne devons vérifier seulement l'injectivité. Supposons que $hk = h'k'$. Alors $e = h^{-1}h'k'k^{-1}$, et donc

$$h^{-1}h' = (k'k^{-1})^{-1} \in H \cap K = \{e\}. \tag{8.54}$$

Par conséquent $|pq| = |H \times K| = |HK|$, et $HK = G$. Le corollaire 3.95 nous indique que

$$G = H \times_\varphi K \tag{8.55}$$

où φ est l'action adjointe. Nous devons maintenant identifier cette action. En d'autres termes, nous savons que $H = \mathbb{Z}/q\mathbb{Z}$ et $K = \mathbb{Z}/p\mathbb{Z}$ et que $\varphi\colon \mathbb{Z}/p\mathbb{Z} \to \mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ est un morphisme. Nous devons déterminer les possibilités pour φ.

Soit n_p le nombre de p-Sylow de G. Comme précédemment, n_p vaut 1, p ou q et la possibilité $n_p = p$ est exclue. Donc n_p est 1 ou q.

Si p ne divise pas $q-1$ Si p ne divise pas $q-1$ alors il n'est pas possible d'avoir $n_p = q$ parce que $n_p \in [1]_p$. Or dire $n_p = q$ demanderait $q \in [1]_p$, c'est à dire $q = kp + 1$, qui impliquerait que p divise $q - 1$.

La seule possibilité est que $n_p = 1$. Dans ce cas, K est également normal dans G. Du coup le produit semi-direct (8.55) est en réalité un produit direct (φ est triviale) et nous avons

$$G = \mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} = \mathbb{Z}/pq\mathbb{Z}. \tag{8.56}$$

Si p divise $q-1$ Cette fois $n_p = 1$ et $n_p = q$ sont tous deux possibles. Ce que nous savons est que $\varphi(\mathbb{Z}/p\mathbb{Z})$ est un sous-groupe de $\mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$. Par le premier théorème d'isomorphisme 3.22, nous avons

$$|\varphi(\mathbb{Z}/p\mathbb{Z})| = \frac{|\mathbb{Z}/p\mathbb{Z}|}{|\ker\varphi|}, \tag{8.57}$$

7. Cette affirmation me semble très bizarre. Comment deux nombres premiers distincts pourraient ne pas être premiers entre eux ? ? ?

ce qui signifie que $|\varphi(\mathbb{Z}/p\mathbb{Z})|$ divise $|\mathbb{Z}/p\mathbb{Z}| = p$. Par conséquent, $|\varphi(\mathbb{Z}/p\mathbb{Z})|$ est égal à 1 ou p. Si c'est 1, alors l'action est triviale et le produit est direct.

Nous supposons que $|\varphi(\mathbb{Z}/p\mathbb{Z})| = p$. Le corollaire 8.21 nous indique que $\text{Aut}(\mathbb{Z}/q\mathbb{Z})$ possède un unique sous-groupe d'ordre p que nous notons Γ ; c'est à dire que $\Gamma = \text{Image}(\varphi)$. Vu que $\varphi\colon \mathbb{Z}/p\mathbb{Z} \to \text{Aut}(\mathbb{Z}/q\mathbb{Z})$ est un morphisme, Γ est généré par $\varphi(\bar{1})$ qui est alors un élément d'ordre p, comme annoncé.

Unicité Nous nous attaquons maintenant à l'unicité. Soient φ et φ' deux morphismes non triviaux $\mathbb{Z}/p\mathbb{Z} \to \text{Aut}(\mathbb{Z}/q\mathbb{Z})$. Étant donné que $\text{Aut}(\mathbb{Z}/q\mathbb{Z})$ ne possède qu'un seul sous-groupe d'ordre p, nous savons que $\text{Image}(\varphi) = \text{Image}(\varphi') = \Gamma$. Nous pouvons donc parler de φ'^{-1} en tant qu'application de $\mathbb{Z}/p\mathbb{Z}$ dans Γ. Nous montrons que

$$\begin{aligned} f\colon \mathbb{Z}/q\mathbb{Z} \times_\varphi \mathbb{Z}/p\mathbb{Z} &\to \mathbb{Z}/q\mathbb{Z} \times_{\varphi'} \mathbb{Z}/p\mathbb{Z} \\ (h,k) &\mapsto (h, \alpha(k)) \end{aligned} \tag{8.58}$$

où $\alpha = \varphi'^{-1} \circ \varphi$ est un isomorphisme de groupes. Le calcul est immédiat :

$$\begin{aligned} f(h_1,k_1)f(h_2mk_2) &= \big(h_1, \alpha(k_1)\big)(h_2, \alpha(k_2)) && \text{(8.59a)}\\ &= \big(h_1\varphi'(\alpha(k_1))h_2m\alpha(k_1k_2)\big) && \text{(8.59b)}\\ &= f\big(h_1\varphi(k_1)h_2, k_1k_2\big) && \text{(8.59c)}\\ &= f\big((h_1,k_1),(h_2,k_2)\big). && \text{(8.59d)} \end{aligned}$$

Par conséquent $\mathbb{Z}/q\mathbb{Z} \times_\varphi \mathbb{Z}/p\mathbb{Z} \simeq \mathbb{Z}/q\mathbb{Z} \times_{\varphi'} \mathbb{Z}/p\mathbb{Z}$.

□

Note : il existe des nombres premiers p et q tels que $q = 1 \mod p$. Par exemple $7 = 1 \mod 3$.

Proposition 8.27 ([12])**.**
Soit G un groupe fini d'ordre pq où p et q sont deux nombres premiers distincts vérifiant

$$\begin{cases} p \neq 1 \mod q & \text{(8.60a)}\\ q \neq 1 \mod p. & \text{(8.60b)} \end{cases}$$

Alors G est cyclique, abélien et

$$G \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}. \tag{8.61}$$

Démonstration. Soient n_p et n_q les nombres de p-Sylow et q-Sylow. Par le théorème de Sylow 8.11, n_p divise pq et $n_p = 1 \mod p$. Le second point empêche n_p de diviser p. Par conséquent n_p divise q et donc n_p vaut 1 ou q. La possibilité $n_p = q$ est exclue par l'hypothèse $q \neq 1 \mod p$. Donc $n_p = 1$, et de la même façon nous obtenons $n_q = 1$.

Soient S l'unique p-Sylow et T, l'unique q-Sylow. Pour les mêmes raisons que celles exposée plus haut, ce sont deux sous-groupes normaux dans G. Étant donné que S est d'ordre p^n pour un certain n et que l'ordre de S doit diviser celui de G, nous avons $|S| = p$. De la même façon, $|T| = q$. Par conséquent S est un groupe cyclique d'ordre p et nous considérons x, un de ses générateurs. De la même façon soit y, un générateur de T.

Nous montrons maintenant que x et y commutent, puis que xy engendre G. Nous savons que $S \cap T$ est un sous-groupe à la fois de S et de T, de telle façon que $|S \cap T|$ divise à la fois $|S| = p$ et $|T| = q$. Nous avons donc $|S \cap T| = 1$ et donc $S \cap T$ se réduit au neutre. Par ailleurs, S et T sont normaux, donc

$$\begin{aligned} (xyx^{-1})y^{-1} &\in T && \text{(8.62a)}\\ x(yx^{-1})y^{-1}) &\in S, && \text{(8.62b)} \end{aligned}$$

donc $xyx^{-1}y^{-1} = e$, ce qui montre que $xy = yx$.

Montrons que xy engendre G. Soit $m > 0$ tel que $(xy)^m = e$. Pour ce m nous avons $x^m = y^{-m}$ et $y^{-m} = x^m$, ce qui signifie que x^m et y^m appartiennent à $S \operatorname{cat} T$ et donc $x^m = y^m = e$. Les

nombres p et q divisent donc tous deux m ; par conséquent $\text{ppcm}(p,q) = pq$ divise m. Nous en concluons que xy est d'ordre pq (il ne peut pas être plus) et qu'il est alors générateur.

Pour la suite nous allons d'abord prouver que $G = ST$ puis que $G \simeq S \times T$. Nous savons déjà que $|S \cap T| = 1$, ce qui nous amène à dire que $|ST| = |S||T|$. En effet si $s, s' \in S$ et $t, t' \in t$ et si $st = s't'$, alors $t = s^{-1}s't'$, ce qui voudrait dire que $s^{-1}s' \in T$ et donc que $s^{-1}s' = e$. Au final nous avons

$$|ST| = |S||T| = pq = |G|. \tag{8.63}$$

Par conséquent $G = ST$. En nous rappelant du fait que $S \cap T = \{e\}$ et que S et T sont normaux, le lemme 3.14 nous dit que $G \simeq S \times T$. Le groupe S étant cyclique d'ordre p nous avons $S = \mathbb{Z}/p\mathbb{Z}$ et pour T, nous avons la même chose : $T = \mathbb{Z}/q\mathbb{Z}$. Nous concluons que

$$G \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}. \tag{8.64}$$

□

Théorème 8.28 (Théorème de Burnside[25])**.**
Le centre d'un p-groupe non trivial est non trivial.

Démonstration. Soit G un p-groupe non trivial. Nous considérons l'action adjointe G sur lui-même. Les points fixes de cette action sont les éléments du centre :

$$\mathcal{Z}_G = \{z \in G \text{ tel que } \sigma_x(z) = z \forall x \in G\} = \text{Fix}_G(G). \tag{8.65}$$

Nous utilisons l'équation aux classes (3.83) pour dire que $|G| = |\mathcal{Z}_G| \mod p$. Mais $|\mathcal{Z}_G|$ n'est pas vide parce qu'il contient l'identité. Donc $|\mathcal{Z}_G|$ est au moins d'ordre p. □

Proposition 8.29.
Si p est un nombre premier, tout groupe d'ordre p ou p^2 est abélien.

Rappel : un groupe d'ordre p ou p^2 est automatiquement un p-groupe.

Démonstration. Si $|G| = p$, alors le théorème de Cauchy 8.2 nous donne l'existence d'un élément d'ordre p. Cet élément est alors automatiquement générateur, G est cyclique et donc abélien.

Si par contre G est d'ordre p^2, alors les choses se compliquent (un peu). D'après le théorème de Burnside 8.28, le centre $\mathcal{Z}$ n'est pas trivial ; il est alors d'ordre p ou p^2. Supposons qu'il soit d'ordre p et prenons $x \in G \backslash \mathcal{Z}$. Alors le stabilisateur de x pour l'action adjointe contient au moins $\mathcal{Z}$ et x, c'est à dire que $|\text{Fix}_G(x)| \geqslant p+1$. Étant donné que $\text{Fix}_G(x)$ est un sous-groupe, son ordre est automatiquement 1, p ou p^2. En l'occurrence, il doit être p^2 (parce que plus grand que p), et donc x doit être central, ce qui est une contradiction. □

8.2.5 Fonction indicatrice d'Euler

Corollaire 8.30.
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,

$$\varphi(pq) = (p-1)(q-1). \tag{8.66}$$

Démonstration. Nous savons que si p et q sont premiers entre eux, alors le théorème 8.26 nous donne l'isomorphisme de groupe

$$(\mathbb{Z}/pq\mathbb{Z}, +) \simeq (\mathbb{Z}/p\mathbb{Z}, +) \times (\mathbb{Z}/q\mathbb{Z}, +). \tag{8.67}$$

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 3.107, 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 $\{1, \ldots, p-1\}$ sont premiers avec p. □

8.3 Le groupe symétrique

La définition des permutations et du groupe symétrique sont 3.90. Voir aussi le thème 56.

8.3.1 Décomposition en cycles

Définition 8.31.
*Le **support** d'une permutation σ est l'ensemble constitué des éléments modifiés par σ :*

$$\operatorname{supp}\sigma = \{i \in \{1,\dots,n\} \text{ tel que } \sigma(i) \neq i\}.$$

Définition 8.32.
*Nous disons qu'un élément $\sigma \in S_n$ **inverse** les nombres $i < j$ si $\sigma(i) > \sigma(j)$. Soit N_σ le nombre d'inversions que $\sigma \in S_n$ possède (c'est le nombre de couples (i,j) avec $i<j$ tels que $\sigma(i) > \sigma(j)$). L'entier*

$$\epsilon(\sigma) = (-1)^{N_\sigma} \tag{8.68}$$

*est la **signature** de σ.*

Un élément du groupe symétrique S_n peut être décomposé en produit de cycles de supports disjoints de la façon suivante. Pour $\sigma \in S_n$, nous écrivons d'abord le cycle qui correspond à l'orbite de 1. Ce sera le cycle

$$(1, \sigma(1), \sigma^2(1), \dots, \sigma^k(1)) \tag{8.69}$$

avec $\sigma^{k+1}(1) = 1$. Ensuite nous recommençons avec le plus petit élément de $\{1,\dots,n\}$ à ne pas être dans ce cycle, et puis le suivant, etc. La *structure* d'une telle décomposition est la donnée des nombres k_i donnant le nombre de cycles de longueur i.

Lemme 8.33 ([22])**.**
Soit $\sigma = (i_1,\dots,i_k) \in S_n$, un cycle de longueur k et $\theta \in S_n$. Alors

$$\theta\sigma\theta^{-1} = \big(\theta(i_1),\dots,\theta(i_k)\big). \tag{8.70}$$

Tous les cycles de longueur k sont conjugués entre eux.

Proposition 8.34 (Classes de conjugaison et structure en cycles[89])**.**
Une classe de conjugaison dans S_n est formée des permutations ayant une décomposition en cycles disjoints de même structure. Autrement dit, deux permutations σ et σ' sont conjuguées si et seulement si le nombre k_i de cycles de longueur i dans σ est le même que le nombre k'_i de cycles de longueur i dans σ'.

Démonstration. Soit $\sigma = c_1 \dots c_m$ la décomposition de σ en cycles de supports disjoints. Les c_i sont des cycles de supports disjoints. Si τ est une permutation, alors

$$\sigma' = \tau\sigma\tau^{-1} = (\tau c_1 \tau^{-1}) \dots (\tau c_m \tau^{-1}), \tag{8.71}$$

mais $\tau c_i \tau^{-1}$ est un cycle de même longueur que c, puisque le lemme 8.33 nous dit que si $\sigma = (a_1,\dots,a_k)$, alors $\tau c \tau^{-1} = \big(\tau(a_1),\dots,\tau(a_k)\big)$. Notons encore que les cycles $\tau c_i \tau^{-1}$ restent à support disjoints.

Donc tous les éléments de la classe de conjugaison de σ sont des permutations de même structure de σ.

Réciproquement, si $\sigma' = c'_1 \dots c'_m$ est une décomposition de σ' en cycles disjoints tels que la longueur de c_i est la même que la longueur de c'_i, alors il suffit de construire des permutations τ_i telles que $\tau_i c_i \tau_i^{-1} = c'_i$, à travers le lemme 8.33. Comme les supports des c_i et des c'_i sont disjoints, la permutation $\tau_1 \dots \tau_m$ conjugue σ et σ'. □

Exemple 8.35
Voyons les classes de conjugaison de S_3. Étant donné que ce groupe agit par définition sur un

ensemble à 3 éléments, aucun élément de S_3 ne possède un un cycle de plus de 3 éléments. Il y a donc seulement des cycles de longueur deux ou trois (à part les triviaux). Aucun élément de S_3 n'a une décomposition en cycles disjoints contenant deux cycles de deux ou un cycle de deux et un de trois.

En résumé il y a trois classes de conjugaison dans S_3. La première est celle contenant seulement l'identité. La seconde est celle contenant les cycles de longueur deux et la troisième contient les cycles de longueur 3.

Ce sont donc

$$C_1 = \{\mathrm{Id}\} \tag{8.72a}$$
$$C_2 = \{(1,2),(1,3),(2,3)\} \tag{8.72b}$$
$$C_3 = \{(1,2,3),(2,1,3)\}. \tag{8.72c}$$

△

Exemple 8.36
Les classes de conjugaison de S_4. Nous savons que les classes de conjugaison dans S_4 sont caractérisées par la structure des décompositions en cycles (proposition 8.34). Le groupe symétrique S_4 possède dont les classes de conjugaison suivantes.

(1) Le cycle vide qui représente la classe constituée de l'identité seule.
(2) Les transpositions (de type (a,b)) qui sont au nombre de 6.
(3) Les 3-cycles. Pour savoir quel est leur nombre nous commençons par remarquer qu'il y a 4 façons de prendre 3 nombres parmi 4 et ensuite 2 façons de les arranger. Il y a donc 8 éléments dans cette classe de conjugaison.
(4) Les 4-cycles. Le premier est arbitraire (parce que c'est cyclique). Pour le second il y a 3 possibilités, et deux possibilités pour le troisième ; le quatrième est alors automatique. Cette classe de conjugaison contient donc 6 éléments.
(5) Les doubles transpositions, du type $(a,b)(c,d)$. Dans ce cas, tous les nombres sont permutés, et l'image de 1 détermine la double transposition. Il y a 3 images possibles, et donc 3 éléments dans cette classe.

△

Lemme 8.37 ([12])**.**
Un k-cycle est une permutation impaire si k est pair et paire si k est impair.

Proposition 8.38.
Tout élément de S_n peut être écrit sous la forme d'un produit fini de transpositions.

Cette décomposition n'est pas à confondre avec celle en cycles de support disjoints. Par exemple $(1,2,3) = (1,3)(1,2)$.

Proposition 8.39 ([22])**.**
Soit S_n le groupe symétrique.

(1) L'application $\epsilon\colon S_n \to \{1,-1\}$ est l'unique homomorphisme surjectif de S_n sur $\{-1,1\}$.
(2) Si $s = t_1 \cdots t_k$ est le produit de k transpositions, alors $\epsilon(s) = (-1)^k$.

Démonstration. Soit $\sigma, \theta \in S_n$. Afin de montrer que $\epsilon(\sigma\theta) = \epsilon(\sigma)\epsilon(\theta)$, nous divisons les couples (i,j) tels que $i \leqslant j$ en 4 groupes suivant que $\theta(i) \gtrless \theta(j)$ et $\sigma\big(\theta(i)\big) \gtrless \sigma\big(\theta(j)\big)$. Nous notons N_1, N_2, N_3 et N_4 le nombre de couples dans chacun des quatre groupes :

(i,j)	$\sigma\big(\theta(i)\big) < \sigma\big(\theta(j)\big)$	$\sigma\big(\theta(i)\big) > \sigma\big(\theta(j)\big)$
$\theta(i) < \theta(j)$	N_1	N_2
$\theta(i) > \theta(j)$	N_3	N_4

Nous avons immédiatement $N_\theta = N_3 + N_4$ et $N_{\sigma\theta} = N_2 + N_4$. Les éléments qui participent à N_σ sont ceux où $\theta(i)$ et $\theta(j)$ sont dans l'ordre inverse de $\sigma\big(\theta(i)\big)$ et $\sigma\big(\theta(j)\big)$ (parce que θ est une bijection). Donc $N_\sigma = N_2 + N_3$. Par conséquent nous avons

$$\epsilon(\sigma)\epsilon(\theta) = (-1)^{N_2+N_3}(-1)^{N_3+N_4} = (-1)^{N_2+N_4} = (-1)^{N_{\sigma\theta}} = \epsilon(\sigma\theta). \tag{8.73}$$

Nous avons prouvé que ϵ est un homomorphisme. Pour montrer que ϵ est surjectif sur $\{-1,1\}$ nous devons trouver un élément $\tau \in S_n$ tel que $\epsilon(\tau) = -1$. Si τ est la transposition $1 \leftrightarrow 2$ alors le couple $(1,2)$ est le seul à être inversé par τ et nous avons $\epsilon(\tau) = -1$.

Avant de montrer l'unicité, nous montrons que si $\sigma = t_1 \dots t_k$ alors $\epsilon(\sigma) = (-1)^k$. Pour cela il faut montrer que $\epsilon(\tau) = -1$ dès que τ est une transposition. Soit τ_{ij}, la transposition (i,j) et $\theta = (i, i+1, \dots, j-1)$ alors le lemme 8.33 dit que

$$\tau_{ij} = \theta\tau_{j-1,j}\theta^{-1}. \tag{8.74}$$

La signature étant un homomorphisme,

$$\epsilon(\tau_{ij}) = \epsilon(\theta)\epsilon(\tau_{j-1,j})\epsilon(\theta)^{-1} = \epsilon(\tau_{j-1,j}) = -1. \tag{8.75}$$

Nous passons maintenant à la partie unicité de la proposition. Soit un homomorphisme surjectif $\varphi\colon S_n \to \{-1,1\}$ et τ, une transposition telle que $\varphi(\tau) = -1$ (qui existe parce que sinon φ ne serait pas surjectif[8]). Si τ' est une autre transposition, il existe $\sigma \in S_n$ tel que $\tau' = \sigma\tau\sigma^{-1}$ (lemme 8.33). Dans ce cas, $\varphi(\tau') = \varphi(\tau) = -1$, et si $\sigma = \tau_1 \dots \tau_k$),

$$\varphi(\sigma) = (-1)^k = \epsilon(\sigma). \tag{8.76}$$

□

8.3.2 Le groupe alterné

Définition 8.40.
*Le groupe A_n des permutations paires est la **groupe alterné**.*

Proposition 8.41.
À propos du groupe alterné dans le groupe symétrique.

(1) Le groupe alterné A_n est un sous-groupe caractéristique[9] *de S_n*

(2) Le sous-groupe A_n est d'indice 2 dans S_n.

(3) Le sous-groupe A_n est l'unique sous-groupe d'indice 2 de S_n.

Démonstration. Soit $\alpha \in \mathrm{Aut}(S_n)$. Étant donné que $\epsilon \circ \alpha$ est un homomorphisme surjectif sur $\{-1,1\}$, par unicité de cet homomorphisme, nous avons $\epsilon \circ \alpha = \epsilon$, et donc $\alpha(A_n) = A_n$. Par le premier théorème d'isomorphisme 3.22, il existe un isomorphisme

$$f\colon S_n/\ker\epsilon \to \mathrm{Image}(\epsilon). \tag{8.77}$$

En égalant le nombre d'éléments nous avons $|S_n : \ker\epsilon| = |S_n : A_n| = 2$.

Nous prouvons maintenant l'unicité. Soit H un sous-groupe d'indice 2 dans S_n. Par le lemme 3.59, H est distingué et nous pouvons considérer le groupe S_n/H. Ce dernier ayant 2 éléments, il est isomorphe à $\{-1,1\}$. Soit θ l'isomorphisme. On note φ le morphisme canonique $\varphi\colon S_n \to S_n/H$:

$$S_n \xrightarrow{\varphi} S_n/H \xrightarrow{\theta} \{-1,1\}. \tag{8.78}$$

La composition $\varphi\circ\theta$ est alors un homomorphisme surjectif de S_n sur $\{-1,1\}$ et nous avons $\varphi\circ\theta = \epsilon$ par la proposition 8.39. L'enchaînement (8.78) nous montre que $H = \ker(\theta\circ\varphi) = \ker(\epsilon) = A_n$. □

8. Nous utilisons ici le fait que tous les éléments de S_n sont des produits de transpositions, proposition 8.38.
9. Définition 3.4.

Proposition 8.42 ([90])**.**
Le groupe symétrique S_n peut être écrit comme un produit semi-direct[10] *du groupe alterné :*

$$S_n = A_n \times_\varphi \mathbb{Z}/2\mathbb{Z} \tag{8.79}$$

où l'action de $\mathbb{Z}/2\mathbb{Z}$ sur A_n est la conjugaison par $\sigma = (12)$, c'est à dire $\rho(-1)\tau = \sigma\tau\sigma^{-1}$.

Démonstration. Nous avons la suite exacte

$$1 \xrightarrow{i} A_n \xrightarrow{i} S_n \xrightarrow{\epsilon} \{\pm 1\} \longrightarrow 1 \tag{8.80}$$

où les i représentent des inclusions et ϵ est la signature définie en 8.32. Grâce à cette suite et au fait que la signature soit un isomorphisme à partir de la partie $\{\mathrm{Id}, \sigma\}$ (pour σ d'ordre 2, par exemple $\sigma = (12)$), le théorème 3.94 nous dit que

$$S_n \simeq A_n \times_\varphi \{\mathrm{Id}, \sigma\} \tag{8.81}$$

où φ est l'action adjointe de $\{\mathrm{Id}, \sigma\}$ sur A_n. □

Lemme 8.43.
Le groupe dérivé du groupe symétrique est le groupe alterné : $D(S_n) = A_n$.

Démonstration. Tout élément de $D(S_n)$ s'écrit sous la forme $ghg^{-1}h^{-1}$. Quel que soit le nombre de transpositions dans g et h, le nombre de transpositions dans $[g, h]$ est pair. □

Proposition 8.44 ([91])**.**
Soit $n \geqslant 3$. Les 3-cycles $c_i = (1, 2, i)$ avec $i = 3, \ldots, n$ engendrent le groupe alterné A_n.

Démonstration. Soit H, le groupe engendré par les c_i. D'abord nous avons

$$c_i = (1, 2, i) = (1, 2)(2, i), \tag{8.82}$$

de telle sorte que $\epsilon(c_i) = 1$. Par conséquent nous avons $H \subset A_n$. Nos montrons par récurrence que $A_n \subset H$.

Pour $n = 3$ il suffit de vérifier que $A_3 = \{\mathrm{Id}, c_3, c_3^2\}$. Supposons avoir obtenu le résultat pour A_{n-1}, et prouvons le pour A_n. Soit $s \in A_n$.

Si $s(n) = n$, alors s se décompose de la même manière que sa restriction s' à $\{1, \ldots, n-1\}$. Par l'hypothèse de récurrence, cette restriction, appartenant à A_{n-1}, se décompose en produit des $c_3, \ldots, c_{n-1}$ et de leurs inverses.

Si $s(n) = k$ alors nous considérons l'élément $c_n^2 c_k s$. Cet élément envoie n sur n et peut donc être décomposé avec les c_i $(i = 1, \ldots, n-1)$ en vertu du point précédent. □

Proposition 8.45.
Lorsque $n \geqslant 5$, tous les 3-cycles de A_n sont conjugués. Autrement dit, la classe de conjugaison d'un 3-cycle est l'ensemble des 3-cycles.

Démonstration. Soient les 3-cycles $\sigma = (i_1, i_2, i_3)$ et $\varphi = (j_1, j_2, j_3)$. Nous considérons une bijection α de $\{1, \ldots, n\}$ telle que $\alpha(i_s) = j_s$. Nous avons immédiatement que $\alpha \in S_n$ et que $\alpha\sigma\alpha^{-1} = \varphi$. Donc les 3-cycles sont conjugués dans S_n. Il reste à prouver qu'ils le sont dans A_n.

Si α est une permutation paire, la preuve est terminée. Si α est impaire, alors nous devons un peu la modifier. Vu que $n \geqslant 5$, nous pouvons prendre s et t, des éléments distincts dans $\{1, \ldots, n\} \backslash \{j_1, j_2, j_3\}$ et poser $\tau = (st)$. Vu que la signature est un homomorphisme et que τ et α sont impairs, l'élément $\tau\alpha$ est pair (lemme et proposition 8.37 et 8.38) et est donc dans A_n. Les supports de τ et φ étant disjoints, ces derniers commutent et nous avons

$$(\tau\alpha)\sigma(\tau\alpha)^{-1} = \tau(\alpha\sigma\alpha^{-1})\tau^{-1} = \tau\varphi\tau^{-1} = \varphi. \tag{8.83}$$

Donc σ et φ sont conjugués par $\tau\alpha$ qui est dans A_n. □

Théorème 8.46 ([12])**.**
Le groupe alterné A_n est simple[11] *pour $n \geqslant 5$.*

10. Définition 3.93.
11. Pas de sous-groupes normaux non triviaux, définition 3.5.

Démonstration. Soit N, un sous-groupe normal de A_n non réduit à l'identité. Étant donné que les 3-cycles engendrent A_n (proposition 8.44) et que tous les 3-cycles sont conjugués dans A_n (proposition 8.45), il suffit de montrer que N contient un 3-cycle. En effet si N contient un 3-cycle, le fait qu'il soit normal implique (par conjugaison) qu'il les contienne tous et donc qu'il contient une partie génératrice de A_n.

Soit donc $\sigma \in N$ différent de l'identité. Nous prenons i dans le support de σ et $j = \sigma(i)$. Nous choisissons ensuite $k \in \{1, \dots, n\}\backslash\{i, j, \sigma^{-1}(i)\}$ et $m = \sigma(k)$. Nous considérons la permutation $\alpha = (ijk)$. Étant donné que N est normal l'élément

$$\theta = (\alpha^{-1}\sigma\alpha)\sigma^{-1} \tag{8.84}$$

est dans N. De plus en utilisant le lemme 8.33 et le fait que $\alpha^{-1} = (ikj)$ nous avons

$$\theta = (ikj)(j\sigma(j)m). \tag{8.85}$$

Cela n'est pas spécialement un 3-cycle, mais nous allons en construire un. Nous allons déterminer que θ est soit un 5-cycle, soit un 3-cycle , soit un 2×2-cycle suivant les valeurs de $\sigma(j)$ et m.

Souvenons-nous que nous avons :

— $i \neq j = \sigma(i)$, puisque i est dans le support de σ ;

— $k \neq i$ et $k \neq j$, par définition de k (rappelons aussi que $k \neq \sigma^{-1}(i)$) ;

— $m \neq i$, $m \neq j$ et $m \neq \sigma(j)$ puisque $m = \sigma(k)$.

Il ne nous reste alors seulement les deux possibilités suivantes :

(1) soit $m = k$, soit $m \neq k$, d'une part ;

(2) soit $\sigma(j) = i$, soit $\sigma(j) = k$, soit $\sigma(j)$ n'est ni i, ni k, ni m, d'autre part.

Supposons dans un premier temps que $m = k$; alors

$$\theta = (ik)(j\sigma(j)). \tag{8.86}$$

C'est a priori un 2×2-cycle. Mais si de plus $\sigma(j) = i$, alors

$$\theta = (ijk) \tag{8.87}$$

qui est un 3-cycle ; et si $\sigma(j) = k$, alors

$$\theta = (ikj) \tag{8.88}$$

qui est un autre 3-cycle.

Supposons à présent que $m \neq k$. Si $\sigma(j)$ n'est ni i, ni k, ni m, alors i, j, k, $\sigma(j)$ et m sont cinq nombres différents, et

$$\theta = (i, j, \sigma(j), m, k) \tag{8.89}$$

est un 5-cycle. Si $\sigma(j) = i$, alors

$$\theta = (ikj)(jim) = (imk) \tag{8.90}$$

qui est un 3-cycle. Si $\sigma(j) = k$, alors

$$\theta = (ikj)(jkm) = (ikm) \tag{8.91}$$

qui est encore un 3-cycle.

Bref nous avons montré que θ est soit un 3-cycle, soit un 5-cycle, soit un 2×2-cycle. Si θ est un 3-cycle, la preuve est terminée.

Si $\theta = (ab)(cd)$, alors on considère $e \in \{1, \dots, n\}\backslash\{a, b, c, d\}$ et nous avons

$$\underbrace{(abe)^{-1}\theta(abe)}_{\in N}\theta^{-1} = (aeb)(ab)(cd)(abe)(an)(cd) = (abe) \in N. \tag{8.92}$$

Si θ est le 5-cycle $(abcde)$, alors l'élément suivant est dans N :

$$(abc)^{-1}\theta(abc)\theta^{-1} = (acb)(abcde)(abc)(aedcb) = (acd). \tag{8.93}$$

Dans tous les cas nous avons trouvé un 3-cycle dans N et nous avons par conséquent $N = A_n$, ce qui fait que A_n ne contient pas de sous-groupes normaux non triviaux. Le groupe alterné A_n est donc simple. □

Nous en déduisons immédiatement que si $n \geqslant 5$, le groupe dérivé de A_n est A_n parce que A_n ne contient pas d'autres sous-groupes non triviaux.

Le théorème suivant montre que tout groupe peut être vu, en agissant sur lui-même, comme une partie du groupe symétrique.

Théorème 8.47.
Un groupe G est isomorphe à un sous-groupe de son groupe symétrique $S(G)$.

Démonstration. Nous considérons φ, la translation à gauche :

$$\begin{aligned} \varphi\colon G &\to S(G) \\ g &\mapsto t_g \end{aligned} \tag{8.94}$$

où $f_g(h) = gh$. Étant donné que

$$\varphi(gh) = ghx = g(t_h x) = t_g \circ t_h(x), \tag{8.95}$$

l'application φ est un morphisme de groupes. Il est injectif parce que si $gx = hx$ pour tout x, en particulier pour $x = e$ nous trouvons $g = h$.

De la même manière, $\varphi(g)x = \varphi(g)y$ implique $x = y$. Cela montre que l'image est bien dans le groupe symétrique.

L'ensemble Image(φ) est donc un sous-groupe de $S(G)$, et φ est un isomorphisme vers ce groupe. □

Lemme 8.48.
Si $n \geqslant 3$, alors

(1) Le centre de S_n est trivial.

(2) Le groupe S_n est non abélien.

Démonstration. Soit $s \in Z(S_n)$ et trois éléments distincts a, b et c de $\{1, \dots, n\}$. Nous posons $\tau = (ab)$ et nous avons $s\tau = \tau s$. En notant $a' = s(a)$ et $b' = s(b)$ nous avons

$$a' = s(a) = (\tau s \tau^{-1})(a) = (\tau s)(b) = \tau(b') \tag{8.96a}$$

$$b' = s(b) = (\tau s \tau^{-1})(b) = (\tau s)(a) = \tau(a'). \tag{8.96b}$$

Donc τ permute a' et b'. Mais comme τ ne permute que a et b, en tant qu'ensembles, $\{a, b\} = \{s(a), s(b)\}$. Le même raisonnement sur $\{b, c\}$ donne $\{b, c\} = \{s(b), s(c)\}$. Et vu que a, b et c sont distincts,

$$\{b\} = \{b, c\} \cap \{a, b\} = \{s(b)\}. \tag{8.97}$$

Cela montre que $s(b) = b$, et donc que le centre de S_n est réduit à la permutation identité.

En ce qui concerne le fait que S_n est non abélien, si nous avions $st = ts$ pour tout $s, t \in S_n$ alors $s = tst^{-1}$ pour tout t. Alors s serait dans le centre de S_n. En bref, si S_n était abélien, son centre serait S_n et non $\{\mathrm{Id}\}$.

□

Proposition 8.49 ([92, 86])**.**
Tout groupe simple[12] *d'ordre* 60 *est isomorphe au groupe alterné A_5.*

Démonstration. Nous avons la décomposition en nombres premiers $60 = 2^2 \cdot 3 \cdot 5$. Déterminons pour commencer le nombre n_5 de 5-Sylow dans G. Le théorème de Sylow 8.11(4) nous renseigne que n_5 doit diviser 60 et doit être égal à $1 \mod 5$. Les deux seules possibilités sont $n_5 = 1$ et $n_5 = 6$. Étant donné que tous les p-Sylow sont conjugués, si $n_5 = 1$ alors le 5-Sylow serait un sous-groupe invariant à l'intérieur de G, ce qui est impossible vu que G est simple. Donc $n_5 = 6$.

12. Définition 3.5.

Par le point (3) du théorème de Sylow, le groupe G agit transitivement sur l'ensemble des 5-Sylow par l'action adjointe :

$$g \cdot S = gSg^{-1}. \tag{8.98}$$

Cela donne donc un morphisme $\theta\colon G \to S_6$. Le noyau de θ est un sous-groupe normal. En effet si $k \in \ker\theta$ et si $g \in G$ nous avons

$$\begin{aligned}(gkg^{-1}) \cdot S &= gkg^{-1}Ggk^{-1}g^{-1} && \text{(8.99a)}\\ &= gkTk^{-1}g^{-1} && \text{(8.99b)}\\ &= gTg^{-1} && \text{(8.99c)}\\ &= S && \text{(8.99d)}\end{aligned}$$

où T est le Sylow $T = g^{-1}Sg$. Étant donné que $k \in \ker\theta$ nous avons utilisé $kTk^{-1} = aT$. Au final $gkg^{-1} \cdot S = S$, ce qui prouve que $gkg^{-1} \in \ker\theta$.

Étant donné que $\ker\theta$ est normal dans G, soit est soit réduit à $\{e\}$ soit il vaut G. La seconde possibilité est exclue parce qu'elle reviendrait à dire que G agit trivialement, ce qui n'est pas correct étant donné qu'il agit transitivement. Nous en déduisons que $\ker\theta = \{e\}$, que θ est injective et que G est isomorphe à un sous-groupe de S_6.

Par ailleurs le groupe dérivé de G est un sous-groupe normal (et non réduit à l'identité parce que G est non commutatif). Donc $D(G) = G$. Étant donné que $G \subset S_6$, nous avons

$$G = D(G) \subset D(S_6) = A_6 \tag{8.100}$$

parce que le groupe dérivé du groupe symétrique est le groupe alterné (lemme 8.43).

L'ensemble $\theta^{-1}(A_6)$ est distingué dans G. En effet si $\sigma \in A_6$ et si $g \in G$ nous avons

$$\theta\big(g\theta^{-1}(\sigma)g^{-1}\big) = \theta(g)\sigma\theta(g)^{-1} \in A_6. \tag{8.101}$$

Nous en déduisons que $\theta^{-1}(A_6)$ est soit G entier soit réduit à $\{e\}$. Si $\theta^{-1}(A_6) = \{e\}$, alors pour tout $g \in G$ nous aurions $g^2 = e$ parce que $\theta(g^2) \in A_6$. L'ordre de G étant 60, il n'est pas possible que tous ses éléments soient d'ordre 2. Nous en déduisons que $\theta(G) \subset A_6$.

Nous nommons $H = \theta(G)$ et nous considérons l'ensemble $X = A_6/H$ où les classes sont prises à gauche, c'est à dire

$$[\sigma] = \{h\sigma \text{ tel que } h \in H\}. \tag{8.102}$$

Évidemment A_6 agit sur X de façon naturelle. Au niveau de la cardinalité,

$$\mathrm{Card}(X) = \frac{|A_6|}{|G|} = \frac{360}{60} = 6. \tag{8.103}$$

Le groupe A_6 agit sur X qui a 6 éléments. Nous avons donc une application $\varphi\colon A_6 \to A_6$. Encore une fois, la simplicité de A_6 montre que $\varphi(A_6) = A_6$.

Nous étudions maintenant $\varphi(H)$ agissant sur X. Un élément $x \in A_6$ fixe la classe de l'unité $[e]$ si et seulement si $x \in H$ et par conséquent $\varphi(H)$ est la fixateur de $[e]$ dans X. À la renumérotation près, nous pouvons identifier $\varphi(H)$ au sous-groupe de A_6 agissant sur $\{1,\dots,6\}$ et fixant 6. Nous avons alors $\varphi(H) = S_5 \cap A_6 = A_5$. Nous venons de prouver que φ fournit un isomorphisme entre A_5 et H. Étant donné que H était isomorphe à G, nous concluons que G est isomorphe à A_6. □

8.3.3 Sous-groupes normaux

8.50 ([93])**.**
Soit le groupe V_4 engendré par les bitranspositions de S_4. Nous savons de l'exemple 8.36(5) que ce groupe contient exactement 3 éléments non triviaux et l'identité. De plus, comme c'est une classe de conjugaison, V_4 est normal dans S_4.

Lemme 8.51.
Les sous-groupes $\mathrm{Fix}_{S_n}(a)$ *(avec* $a \in \{1,\dots,n\}$*) sont conjugués entre eux.*

Démonstration. Soit $\sigma \in \mathrm{Fix}(a)$ et $s \in S_n$ nous devons prouver que $s\sigma s^{-1}$ est le fixateur d'un élément de $\{1,\dots,n\}$. Nous notons $s(a) = b$. Alors

$$(s\sigma s^{-1})(b) = (s\sigma)(a) = s(a) = b. \tag{8.104}$$

Donc $s\,\mathrm{Fix}(a)s^{-1} \subset \mathrm{Fix}(b)$.

Dans l'autre sens, si $\sigma \in \mathrm{Fix}(b)$ alors $s^{-1}\sigma s \in \mathrm{Fix}(a)$. Mais $\sigma = s(s^{-1}\sigma s)s^{-1}$, donc $\sigma \in s\,\mathrm{Fix}(a)s^{-1}$. □

Proposition 8.52 (Sous-groupes normaux de S_n [93])**.**
Les sous-groupes normaux de S_n ne sont pas légions.

(1) Pour $n = 4$, le sous-groupes normaux de S_4 sont $\{\mathrm{Id}\}$, V_4, A_4 *et* S_4.

(2) Pour $n \neq 4$, les sous-groupes normaux de S_n sont $\{\mathrm{Id}\}$, A_n *et* S_n.

Démonstration. Les cas $n \leqslant 2$ sont un peu triviaux, donc nous faisons $n \geqslant 3$. Soit H normal dans S_n et $s \neq \mathrm{Id}$ dans H ; par le lemme 8.48, s n'est pas dans le centre de S_n et il existe $u \in S_n$ tel que $us \neq su$. Vu que u est un produit de transpositions (proposition 8.38), il existe une transposition t telle que $st \neq ts$. Le sous-groupe H est normal et que $s \in H$ nous avons aussi $ts^{-1}t^{-1} \in H$. Mais en même temps, la combinaison sts^{-1} est le conjugué d'une transposition et est donc également une transposition (classe de conjugaison de S_4 dans 8.36). Nous en concluons que $sts^{-1}t^{-1}$ est un produit de deux transpositions appartenant H.

Nous venons de prouver que H contient au moins un produit de deux transpositions. Et ce produit est différent de Id parce que $sts^{-1}t^{-1} = \mathrm{Id}$ impliquerait $st = ts$.

Soient donc deux transpositions $t_1, t_2 \in H$ telles que $t_1t_2 \neq \mathrm{Id}$. Les supports de t_1 et t_2 ont soit 1 soit aucun éléments communs.

Premier cas Supposons $t_1 = (a,b)$, $t_2 = (b,c)$ avec a,b,c distincts dans $\{1,\dots,n\}$. Dans ce cas $t_1t_2 = (a,b,c)$ et H contient un cycle de longueur 3. Vu que H est normal et que les cycles de longueur trois sont une classe de conjugaison (exemple 8.36) et que A_n est engendré par ceux-ci (proposition 8.44), $A_n \subset H$. Mais A_n est d'indice deux dans S_n (proposition (2)(2)). Quel nombre plus grand que $n!/2$ divise $n!$? Seulement n lui-même. Donc H est soit A_n soit S_n.

Second cas Le groupe H contient un élément de la forme $(ab)(cd)$ avec a,b,c,d distincts dans $\{1,\dots,n\}$.

Si $n = 3$ Impossible parce que avec $n = 3$ nous n'avons pas quatre éléments distincts.

Si $n = 4$ Le sous-groupe H de S_4 contient un élément de V_4 qui n'est pas l'identité. Par normalité et classes de conjugaison, H contient V_4. Nous devons maintenant prouver que si H n'est pas V_4 alors H est A_4 ou S_4. Nous avons les inclusions $V_4 \subset H \subset S_4$ et donc les inégalités

$$4 \leqslant |H| \leqslant 24. \tag{8.105}$$

Donc le nombre $|H|$ est un multiple de 4 qui divise 24. Les possibilités sont $|H| = 4, 8, 12, 24$. La possibilité $|H| = 4$ donne $H = V_4$; si $|H| = 24$ alors $H = S_4$; si $|H| = 12$ alors H est d'indice 2 dans S_4 et $H = A_n$ (proposition 8.41(3)). Quid de $|H| = 8$?

D'après le corollaire 3.55 au théorème de Lagrange, l'ordre d'un élément divise l'ordre du groupe. Soit x dans H mais pas dans V_4. L'ordre de x peut être 1, 2, 4 ou 8. Ordre 1 serait $x = \mathrm{Id}$. Ordre 8, pas possible parce que S_4 n'a pas d'éléments d'ordre 8.

x d'ordre 2 Prenons la décomposition de x en cycles disjoints. Vu qu'on est dans S_4, ces cycles ne peuvent être que des transpositions. Soit il y en a un (alors H contient une transposition et donc $H = S_4$), soit il y en a deux et alors x est dans V_4.

x d'ordre 4 L'élément x serait alors un cycle de longueur 4. Et alors H contient tous les cycles de longueur 4. Par exemple il contient le produit $(abcd)(bacd) = (adc)$. Le sous-groupe H contient alors A_4 (parce qu'il contient tous les 3-cycles).

Si $n \geqslant 5$ Soit un élément e distinct de a, b, c et d. Par notre liste préférée des classes de conjugaisons (exemple 8.36(5)), le 2-cycle $(c, e)(a, b)$ est conjugué à $(a, b)(c, d)$ et appartient donc à H. Mais alors le produit suivant est également dans H :

$$(ce)(ab)(ab)(cd) = (ce)(cd) = (ecd). \tag{8.106}$$

Donc H contient un 3-cycle, et par conséquent tous les 3-cycles. Encore une fois, cela prouve que H est soit A_n soit S_n.

Pourquoi $n = 4$ est spécial ? Dans le premier cas, nous montrons tout de suite que $H = V_4$ n'est pas possible. Dans le deuxième cas, nous montrons que, grâce à un élément différent de a, b, c et d, la possibilité $H = V_4$ est exclue. La possibilité $H = V_4$ n'existe que pour $n = 4$.

□

8.3.4 Indice

Théorème 8.53.
Tout sous-groupe d'indice n dans S_n est isomorphe à S_n.

Démonstration. Pour $n = 1$, il n'y a rien. Pour $n = 2$, un sous-groupe d'indice 2 ne peut contenir que 1 élément qui est donc l'identité. Ok pour que $\{\mathrm{Id}\}$ soit égal à S_1 ?

Pour les autres, il y a un peu plus de travail.

Pour $n = 3$ Nous avons $|S_3| = 6$. Donc un sous-groupe d'indice 3 dans S_3 contient exactement 2 éléments. Il contient Id et un autre élément $\sigma \in S_3$ qui doit vérifier $\sigma^2 = \mathrm{Id}$ ou $\sigma^2 = \sigma$. Aucun élément de S_3 ne vérifie $\sigma^2 = \sigma$ (à part l'identité). Donc $\sigma^2 = \mathrm{Id}$, ce qui fait que σ est une transposition. Donc

$$H = \{\mathrm{Id}, (12)\} \tag{8.107}$$

ou l'identité avec (23) ou avec (13). Dans tous les cas c'est isomorphe à S_2.

Pour $n = 4$ Nous avons $|S_4 : H| = 4$, donc $|H| = 6$. Mais $6 = 2 \times 3$ et $2 \mid 3 - 1$, donc le théorème 8.26 nous dit que H est soit cyclique[13] (et donc abélien), soit un produit semi-direct. Vu que S_4 n'a pas d'éléments d'ordre 6, aucun sous-groupe d'ordre 6 ne peut être cyclique. Nous sommes donc dans le cas du produit semi-direct

$$H = \mathbb{Z}_3 \times_\varphi \mathbb{Z}_2 \tag{8.108}$$

où $\varphi\colon \mathbb{Z}_2 \to \mathrm{Aut}(\mathbb{Z}_3)$ et $\varphi(1)$ est d'ordre 2 dans $\mathrm{Aut}(\mathbb{Z}_3)$. Il convient de nous attarder un peu pour être sûr d'avoir bien compris tout ce qui se trouve dans l'identification (8.108). D'abord un point de notations : ici nous considérons les groupes $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ munis de l'addition. Donc 1 n'est pas le neutre. Ensuite nous savons du théorème 8.19 que $\mathrm{Aut}(\mathbb{Z}/3\mathbb{Z}) = (\mathbb{Z}/3\mathbb{Z})^*$, et que via cette identification, $\varphi(1) = 2 \in (\mathbb{Z}/3\mathbb{Z})^*$ au sens où $\varphi(1)x = 2x$. Nous avons alors $\varphi(1)^2 x = 4x = x$ dans $\mathbb{Z}/3\mathbb{Z}$. Cela montre bien que $\varphi(1)$ est d'ordre 2.

Par rapport à la proposition 8.42, ici nous écrivons $\mathbb{Z}_2 = \big(\{0, 1\}, +\big)$ alors que là nous écrivons $\mathbb{Z}_2 = \big(\{-1, 1\}, \cdot\big)$. Ce sont les mêmes groupes, mais il convient de remarquer que le 1 ici est le -1 là.

Nous savons par la proposition 8.42 que $S_n = A_n \times_\varphi \mathbb{Z}_2$; en comparant avec (8.108) nous voyons qu'il suffit de prouver que $A_3 = \mathbb{Z}/3\mathbb{Z}$ pour avoir $H = S_3$.

Le groupe A_3 possède $|S_3|/2 = 3$ éléments. Il est vite vu que $A_3 = \{\mathrm{Id}, (12)(31), (12)(32)\}$: ce sont trois éléments de signature paire dans S_3 ; donc c'est S_3. La correspondance $\mathrm{Id} \mapsto 0$, $(12)(13) \mapsto 1$, $(13)(12) \mapsto 2$ donne un isomorphisme avec $(\mathbb{Z}_3, +)$.

Pour $n \geqslant 5$ Soit un sous-groupe H d'indice n dans S_n et l'action à gauche de S_n sur $E = S_n/H$ (qui n'est a priori pas un groupe) donnée par $g \cdot [s] = [gs]$.

13. Définition 3.7.

Morphisme $\varphi\colon S_n \to S_E$ Le φ définit par l'action est un morphisme parce que

$$\varphi(g_1g_2)[s] = [g_1g_2s] = \varphi(g_1)[g_2s] = \varphi(g_1)\varphi(g_2)[s]. \tag{8.109}$$

Mais il faut également vérifier que pour chaque $g \in G$, l'application $\varphi(g)\colon E \to E$ est bien une permutation. Pour l'injectivité, si $\varphi(g)[s_1] = \varphi(g)[s_2]$ alors $[gs_1] = [gs_2]$, donc il existe $h \in H$ tel que $gs_1 = gs_2h$, ce qui prouve que $s_1 = s_2h$ et donc que $[s_1] = [s_2]$. Pour la surjectivité, soit $[t] \in S_n/H$ et résolvons $\varphi(g)[s] = [t]$ par rapport à s. L'élément $s = g^{-1}t$ fonctionne.

$\ker(\varphi)$ est normal Soit $z \in \ker(\varphi)$, c'est à dire que $\varphi(z) = \mathrm{Id}_E$. Alors pour $\sigma \in S_n$ nous avons $\varphi(\sigma z\sigma^{-1}) = \varphi(\sigma)\varphi(z)\varphi(\sigma^{-1}) = \mathrm{Id}_E$.

$\ker(\varphi) = \bigcap_{g\in S_n} gHg^{-1}$ Supposons que $z \in gHg^{-1}$ pour tout g, et calculons $\varphi(z)[s]$. D'abord par hypothèse il existe $h \in H$ tel que $z = shs^{-1}$, donc

$$\varphi(z)[s] = [zs] = [zhs^{-1}s] = [s], \tag{8.110}$$

ce qui prouve que $\varphi(z) = \mathrm{Id}$.

Dans l'autre sens, soit $z \in \ker(\varphi)$. Donc $\varphi(z)[s] = [s]$. Il existe donc $h \in H$ tel que $zs = sh$, c'est à dire tel que $z = shs^{-1}$. La formule demandée est donc prouvée.

Questions d'ordre Nous savons que $|H| = (n-1)!$ alors que $|A_n| = \frac{n!}{2}$. Donc $|H| < |A_n|$ avec une inégalité stricte. En même temps nous avons $|\ker(\varphi)| \leqslant |H|$ parce que $\ker(\varphi)$ est une intersection dont un des termes est H lui-même. Nous avons alors les inégalités

$$|\ker(\varphi)| \leqslant |H| = (n-1)! < |A_n|. \tag{8.111}$$

Mais le seul sous-groupes normaux de S_n sont A_n et S_n et $\{\mathrm{Id}\}$ (proposition 8.52). Donc $\ker(\varphi) = \mathrm{Id}$ et φ est une injection entre deux ensembles finis de même cardinalité. Cela fait de φ une bijection et donc un isomorphisme de groupes

$$\varphi\colon S_n \to S_E. \tag{8.112}$$

Soit une fonction de numérotation $\psi\colon E \to \{1,\dots,n\}$. Avec cela nous définissons un isomorphisme de groupes

$$\begin{aligned}\tilde{\psi}\colon S_E &\to S_n\\ \sigma &\mapsto \psi\sigma\psi^{-1}.\end{aligned} \tag{8.113}$$

Fixilisateur Nous montrons à présent que $(\tilde{\psi}\circ\varphi)(H) = \mathrm{Fix}\big(\psi[\mathrm{Id}]\big)$ où le stabilisateur est pris dans S_n. Pour la première inclusion, soit $h \in H$. Nous avons $(\tilde{\psi}\circ\varphi)(h) = \psi\circ\varphi(h)\psi^{-1}$, qui nous appliquons à $\psi[\mathrm{Id}]$:

$$(\tilde{\psi}\circ\varphi)(h)\psi[\mathrm{Id}] = \psi\circ\varphi(h)[\mathrm{Id}] = \psi[h] = \psi[\mathrm{Id}]. \tag{8.114}$$

Donc $(\tilde{\psi}\circ\varphi)(H) \subset \mathrm{Fix}\big(\psi[\mathrm{Id}]\big)$.

Pour l'autre inclusion, soit $\sigma \in S_n$ tel que $\sigma\psi[\mathrm{Id}] = \psi[\mathrm{Id}]$. Vu que $\sigma \in S_n$ nous avons $s \in S_E$ tel que $\sigma = \tilde{\psi}(s)$. Pour ce s nous avons donc

$$\big(\tilde{\psi}(s)\circ\psi\big)[\mathrm{Id}] = \psi[\mathrm{Id}], \tag{8.115}$$

d'où nous déduisons $s[\mathrm{Id}] = [\mathrm{Id}]$. Cela prouve que s stabilise $[\mathrm{Id}]$ dans S_E. Donc $s = \varphi(h)$ pour un certain $h \in H$, et au final $\sigma = \tilde{\psi}\big(\varphi(h)\big)$.

Conclusion L'application $\tilde{\psi}\circ\varphi\colon H \to S_n$ est une application dont l'image est le fixateur d'un point. Plus précisément,

$$\tilde{\psi}\circ\varphi\colon H \to \mathrm{Fix}\big(\psi[\mathrm{Id}]\big) \tag{8.116}$$

est un isomorphisme de groupe. Mais le stabilisateur d'un point dans S_n est S_{n-1}.

□

Chapitre 9

Suite des espaces vectoriels

9.1 Déterminants

9.1.1 Formes multilinéaires alternées

Définition 9.1.
Soit E, un $\mathbb{K}$-espace vectoriel. Une forme linéaire ***alternée*** *sur E est une application linéaire $f\colon E \to \mathbb{K}$ telle que $f(v_1, \dots, v_k) = 0$ dès que $v_i = v_j$ pour certains $i \neq j$.*

Lemme 9.2.
Une forme linéaire alternée est antisymétrique. Si $\mathbb{K}$ est de caractéristique différente de 2, alors une forme antisymétrique est alternée.

Démonstration. Soit f une forme alternée ; quitte à fixer toutes les autres variables, nous pouvons travailler avec une 2-forme et simplement montrer que $f(x, y) = -f(y, x)$. Pour ce faire nous écrivons

$$0 = f(x + y, x + y) = f(x, x) + f(x, y) + f(y, x) + f(y, y) = f(x, y) + f(y, x). \tag{9.1}$$

Pour la réciproque, si f est antisymétrique, alors $f(x, x) = -f(x, x)$. Cela montre que $f(x, x) = 0$ lorsque $\mathbb{K}$ est de caractéristique différente de deux. □

Proposition 9.3 ([94])**.**
Soit E, un $\mathbb{K}$-espace vectoriel de dimension n, où la caractéristique de $\mathbb{K}$ n'est pas deux. L'espace des n-formes multilinéaires alternées sur E est de $\mathbb{K}$-dimension 1.

Démonstration. Soit $\{e_i\}$, une base de E et $f\colon E \to \mathbb{K}$ une n-forme linéaire alternée, puis $(v_1, \dots, v_n)$ des vecteurs de E. Nous pouvons les écrire dans la base

$$v_j = \sum_{i=1}^{n} \alpha_{ij} e_i \tag{9.2}$$

et alors exprimer f par

$$f(v_1, \dots, v_n) = f\Big(\sum_{i_1=1}^{n} \alpha_{1i_1} e_{i_1}, \dots, \sum_{i_n=1}^{n} \alpha_{ni_n} e_{i_n} \Big) \tag{9.3a}$$

$$= \sum_{i,j} \alpha_{1i_1} \dots \alpha_{ni_n} f(e_{i_1}, \dots, e_{i_n}). \tag{9.3b}$$

Étant donné que f est alternée, les seuls termes de la somme sont ceux dont les i_k sont tous différents, c'est à dire ceux où $\{i_1, \dots, i_n\} = \{1, \dots, n\}$. Il y a donc un terme par élément du groupe des permutations S_n et

$$f(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \alpha_{\sigma(1)1} \dots \alpha_{\sigma(n)n} f(e_{\sigma(1)}, \dots, e_{\sigma(n)}). \tag{9.4}$$

En utilisant encore une fois le fait que la forme f soit alternée, $f = f(e_1, \dots, e_n)\Pi$ où

$$\Pi(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) \alpha_{\sigma(1)1} \dots \alpha_{\sigma(n)n}. \tag{9.5}$$

Pour rappel, la donnée des v_i est dans les nombres α_{ij}.

L'espace des n-formes alternées est donc *au plus* de dimension 1. Pour montrer qu'il est exactement de dimension 1, il faut et suffit de prouver que Π est alternée. Par le lemme 9.2, il suffit de prouver que cette forme est antisymétrique [1].

Soient donc $v_1, \dots, v_n$ tels que $v_i = v_j$. En posant $\tau = (1i)$ et $\tau' = (2j)$ et en sommant sur $\sigma\tau\tau'$ au lieu de σ, nous pouvons supposer que $i = 1$ et $j = 2$. Montrons que $\Pi(v, v, v_3, \dots, v_n) = 0$ en tenant compte que $\alpha_{i1} = \alpha_{i2}$:

$$\Pi(v, v, v_3, \dots, v_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) \alpha_{\sigma(1)1} \alpha_{\sigma(2)2} \alpha_{\sigma(3)3} \dots \alpha_{\sigma(n)n} \tag{9.6a}$$

$$= \sum_{\sigma \in S_n} \epsilon(\sigma\tau) \alpha_{\sigma\tau(1)1} \alpha_{\sigma\tau(2)2} \alpha_{\sigma\tau(3)3} \dots \alpha_{\sigma\tau(n)n} \qquad \text{où } \tau = (12) \tag{9.6b}$$

$$= -\sum_{\sigma \in S_n} \epsilon(\sigma) \alpha_{\sigma(1)1} \alpha_{\sigma(2)2} \alpha_{\sigma(3)3} \dots \alpha_{\sigma(n)n} \tag{9.6c}$$

$$= -\Pi(v, v, v_3, \dots, v_n). \tag{9.6d}$$

□

Lemme 9.4.
Soit $\mathbb{K}$ un corps fini autre que $\mathbb{F}_2$ [2] , 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

$$\tilde{\varphi}\colon \frac{\mathrm{GL}(n, \mathbb{K})}{\mathrm{SL}(n, \mathbb{K})} \to M \tag{9.7}$$

vérifiant $\varphi = \tilde{\varphi} \circ \pi$ où π est la projection.

Nous pouvons faire la même chose avec l'application

$$\det\colon \mathrm{GL}(n, \mathbb{K}) \to \mathbb{K}^* \tag{9.8}$$

qui est un morphisme de groupes dont le noyau est $\mathrm{SL}(n, \mathbb{K})$. Cela nous donne une application

$$\tilde{\det}\colon \frac{\mathrm{GL}(n, \mathbb{K})}{\mathrm{SL}(n, \mathbb{K})} \to \mathbb{K}^* \tag{9.9}$$

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

$$\varphi = \tilde{\varphi} \circ \tilde{\det}^{-1} \circ \tilde{\det} \circ \pi. \tag{9.10}$$

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 $\delta' = \delta$. □

1. C'est ici que joue l'hypothèse sur la caractéristique de $\mathbb{K}$.
2. Je ne comprends pas très bien à quel moment joue cette hypothèse.

Théorème 9.5.
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

$$\epsilon(u) = \left(\frac{\det(u)}{p}\right). \tag{9.11}$$

Ici ϵ est la signature de u vue comme une permutation des éléments de $\mathbb{F}_p$.

Démonstration. Commençons par prouver que

$$\epsilon\colon \mathrm{GL}(V) \to \{-1, 1\}. \tag{9.12}$$

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 8.39),

$$\epsilon(uv) = \epsilon(\bar{u} \circ \bar{v}) = \epsilon(\bar{u})\epsilon(\bar{v}). \tag{9.13}$$

Par ailleurs $\{-1, 1\}$ est abélien, donc le lemme 9.4 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 5.140 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

$$\delta(x) = (\delta \circ \det)(u) = \epsilon(u), \tag{9.14}$$

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 5.159(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

$$\begin{aligned} \beta\colon \mathbb{F}_q &\to \mathbb{F}_q \\ x &\mapsto yx. \end{aligned} \tag{9.15}$$

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

$$(1, y, y^2, \dots, y^{q-2}). \tag{9.16}$$

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. □

9.1.2 Déterminant d'une famille de vecteurs

Nous considérons un corps $\mathbb{K}$ et l'espace vectoriel E de dimension n sur $\mathbb{K}$.

Définition 9.6 (Déterminant d'une famille de vecteurs[7])**.**
*Le **déterminant** $(v_1, \dots, v_n)$ dans la base B est l'élément de $\mathbb{K}$*

$$\det_{(e_1,\dots,e_n)}(v_1, \dots, v_n) = \sum_{\sigma \in S_n} \epsilon(\sigma) \prod_{i=1}^{n} e^*_{\sigma(i)}(v_i) \tag{9.17}$$

où la somme porte sur le groupe symétrique, $\epsilon(\sigma)$ est la signature de la permutation σ et e_k^* est le dual de e_k.

Nous le notons $\det_{(e_1,\dots,e_n)}(v_1, \dots, v_n)$.

Lemme 9.7 ([7])**.**
Les propriétés du déterminant. Soit B une base de E.

(1) *L'application* $\det_B : E^n \to \mathbb{K}$ *est n-linéaire et alternée.*
(2) *Pour toute base,* $\det_B(B) = 1$.
(3) *Le déterminant ne change pas si on remplace un vecteur par une combinaison linéaire des autres :*
$$\det_B(v_1, \dots, v_n) = \det_B\big(v_1 + \sum_{s=2}^{n} a_s v_s, v_2, \dots, v_n\big). \tag{9.18}$$
(4) *Si on permute les vecteurs,*
$$\det_B(v_1, \dots, v_n) = \epsilon(\sigma) \det_B(v_{\sigma(1)}, \dots, v_{\sigma(n)}). \tag{9.19}$$
(5) *Si* B' *est une autre base :*
$$\det_B = \det_B(B') \det_{B'} \tag{9.20}$$
(6) *Nous avons aussi la formule* $\det_B(B') \det_{B'}(B) = 1$.
(7) *Les vecteurs* $\{v_1, \dots, v_n\}$ *forment une base si et seulement si* $\det_B(v_1, \dots, v_n) \neq 0$.

Démonstration. Point par point.

(1) En posant $v_1 = x_1 + \lambda x_2$ nous avons
$$\det_B(x_1 + \lambda x_2, v_2, \dots, v_n) = \sum_\sigma \epsilon(\sigma) \prod_{i=1}^{n} e^*_{\sigma(i)}(v_i) \tag{9.21a}$$
$$= \sum_\sigma \epsilon(\sigma) \Big(e^*_{\sigma(1)}(x_1 + \lambda x_2) \Big) \prod_{i=2}^{n} e^*_{\sigma(i)}(v_i). \tag{9.21b}$$
À partir de là, la linéarité de $e^*_{\sigma(1)}$ montre que $\det_B$ est linéaire en son premier argument. Pour les autres argument, le même calcul tient.
En ce qui concerne le fait d'être alternée, permuter v_k et v_l revient à calculer $\det_B(v_{\sigma_{kl}(1)}, \dots, v_{\sigma_{kl}(n)})$, c'est à dire changer la somme $\sum_\sigma$ en $\sum_{\sigma\circ\sigma_{kl}}$. Cela ajoute 1 à $\epsilon(\sigma)$ vu que l'on ajoute une permutation.

(2) Nous avons
$$\det_B(B) = \sum_{\sigma\in S_n} \epsilon(\sigma) \prod_{i=1}^{n} \underbrace{e^*_{\sigma(i)}(e_i)}_{=\delta_{\sigma(i),i}}. \tag{9.22}$$
Si σ n'est pas l'identité, le produit contient forcément un facteur nul. Il ne reste de la somme que $\sigma = \mathrm{Id}$ et le résultat est 1.

(3) Vu que $\det_B$ est linéaire en tous ses arguments,
$$\det_B\big(v_1 + \sum_{s=2}^{n} a_s v_s, v_2, \dots, v_n\big) = \det_B(v_1, \dots, v_n) + \sum_{s=2}^{n} a_s \det_B(v_s, v_2, \dots, v_n). \tag{9.23}$$
Chacun des termes de la somme est nul parce qu'il y a répétition de v_s parmi les arguments alors que la forme est alternée.

(4) Nous devons calculer $\det_B(v_{\sigma(1)}, \dots, v_{\sigma(n)})$, et pour y voir plus clair nous posons $w_i = v_{\sigma(i)}$. Alors :
$$\det_B(v_{\sigma(1)}, \dots, v_{\sigma(n)}) = \sum_{\sigma'} \epsilon(\sigma') \prod_{i=1}^{n} e^*_{\sigma'(i)}(w_i) \tag{9.24a}$$
$$= \sum_{\sigma'} \epsilon(\sigma') \prod_{i=1}^{n} e^*_{\sigma'(i)}(v_{\sigma(i)}) \tag{9.24b}$$
$$= \sum_{\sigma'} \epsilon(\sigma') \prod_{i=1}^{n} e^*_{\sigma^{-1}\sigma'(i)}(v_i) \tag{9.24c}$$
$$= \sum_{\sigma'} \epsilon(\sigma\sigma') \prod_{i=1}^{n} e^*_{\sigma'(i)}(v_i) \tag{9.24d}$$
$$= \epsilon(\sigma) \det_B(v_1, \dots, v_n). \tag{9.24e}$$

Justifications : nous avons d'abord modifié l'ordre des éléments du produit et ensuite l'ordre des éléments de la somme. Nous avons ensuite utilisé le fait que $\epsilon\colon S_n \to \{0,1\}$ était un morphisme de groupe (proposition 8.39).

(5) Étant donné que l'espace des formes multilinéaires alternées est de dimension 1, il existe un $\lambda \in \mathbb{K}$ tel que $\det_B = \lambda \det_{B'}$. Appliquons cela à B' :

$$\det_B(B') = \lambda \det_{B'}(B'), \tag{9.25}$$

donc $\lambda = \det_B(B')$.

(6) Il suffit d'appliquer l'égalité précédente à B en nous souvenant que $\det_B(B) = 1$.

(7) Si $B' = \{v_1, \dots, v_n\}$ est une base alors $\det_B(B') \neq 0$, sinon il n'est pas possible d'avoir $\det_B(B') \det_{B'}(B) = 1$.
À l'inverse, si B' n'est pas une base, c'est que $\{v_1, \dots, v_n\}$ est liée par le théorème 7.14(3). Il y a donc moyen de remplacer un des vecteurs par une combinaison linéaire des autres. Le déterminant s'annule alors.

□

D'après la proposition 9.3, il existe une unique forme n-linéaire alternée égale à 1 sur B, et c'est $\det_B\colon E^n \to \mathbb{K}$.

9.1.3 Déterminant d'un endomorphisme

L'interprétation géométrique du déterminant en termes d'aires et de volumes est donnée après la théorème 13.223.

Lemme 9.8.
Si $f\colon E \to E$ est un endomorphisme, si B et B' sont deux bases, alors $\det_B\big(f(B)\big) = \det_{B'}\big(f(B')\big)$.

Démonstration. L'application

$$\begin{aligned} \varphi\colon E^n &\to \mathbb{K} \\ v_1, \dots, v_n &\mapsto \det_B\big(f(v_1), \dots, f(v_n)\big) \end{aligned} \tag{9.26}$$

est n-linéaire et alternée ; il existe donc $\lambda \in \mathbb{K}$ tel que $\varphi = \lambda \det_B$. En appliquant cela à B :

$$\det_B\big(f(B)\big) = \lambda \det_B(B) = \lambda. \tag{9.27}$$

Nous avons donc déjà prouvé que $\lambda = \det_B\big(f(B)\big)$, c'est à dire

$$\det_B\big(f(v)\big) = \det_B\big(f(B)\big) \det_B(v). \tag{9.28}$$

Nous allons maintenant introduire B' là où il y a du v en utilisant les formules (9.20) :

$$\det_B\big(f(v)\big) = \det_B(B') \det_{B'}\big(f(v)\big) \tag{9.29a}$$

$$\det_B(v) = \det_B(B') \det_{B'}(v). \tag{9.29b}$$

Nous obtenons

$$\det_{B'}\big(f(v)\big) = \det_B\big(f(B)\big) \det_{B'}(v). \tag{9.30}$$

Et on applique cela à $v = B'$:

$$\det_{B'}\big(f(B')\big) = \det_B\big(f(B)\big) \underbrace{\det_{B'}(B')}_{=1}. \tag{9.31}$$

□

Cette proposition nous permet de définir le déterminant d'un endomorphisme de la façon suivante sans préciser la base.

Définition 9.9 ([7]).
*Si $f\colon E \to E$ est un endomorphisme, le **déterminant** de f est*

$$\det(f) = \det_B\big(f(B)\big) \tag{9.32}$$

Couplé à la formule 9.17, nous pouvons écrire la formule pratique à utiliser le plus souvent. Si $\{e_i\}_{i=1,\dots,n}$ est une base orthonormée de E et si $f\colon E \to E$ est un endomorphisme,

$$\det(f) = \sum_{\sigma\in S_n} \epsilon(\epsilon) \prod_{i=1}^{n} \langle e_{\sigma(i)}, f(e_i)\rangle. \tag{9.33}$$

Et si vous avez tout suivi, vous aurez remarqué que les produits scalaires impliqués dans cette formule sont les éléments de la matrice de f dans la base $\{e_i\}$ parce que $\langle e_i, f(e_j)\rangle$ est la composante i de l'image de e_j par f. Si la matrice est composée en mettant en colonne les images des vecteurs de base, le compte est bon.

Proposition 9.10.
Principales propriétés géométriques du déterminant d'un endomorphisme.

(1) Si f et g sont des endomorphismes, alors $\det(f\circ g) = \det(f)\det(g)$.

(2) L'endomorphisme f est un automorphisme[3] si et seulement si $\det(f) \neq 0$.

(3) Si $\det(f) \neq 0$ alors $\det(f^{-1}) = \det(f)^{-1}$.

(4) L'application $\det\colon \mathrm{GL}(E) \to \mathbb{K}\backslash\{0\}$ est un morphisme de groupe.

Démonstration. Point par point.

(1) Nous considérons l'application

$$\begin{aligned}\varphi\colon E^n &\to \mathbb{K}\\ v &\mapsto \det_B\big(f(v)\big).\end{aligned} \tag{9.34}$$

Comme d'habitude nous avons $\varphi(v) = \lambda\det_B(v)$. En appliquant à B et en nous souvenant que $\det_B(B) = 1$ nous avons $\det_B\big(f(B)\big) = \lambda$. Autrement dit :

$$\lambda = \det(f). \tag{9.35}$$

Calculons à présent $\varphi\big(g(B)\big)$: d'une part,

$$\varphi\big(g(B)\big) = \det_B\big((f\circ g)(B)\big) \tag{9.36}$$

et d'autre part,

$$\varphi\big(g(B)\big) = \lambda\det_B\big(g(B)\big) = \lambda\det(g) \tag{9.37}$$

En égalisant et en reprenant la la valeur déjà trouvée de λ,

$$\det\big(f\circ g)(B)\big) = \det(f)\det(g), \tag{9.38}$$

ce qu'il fallait.

(2) Supposons que f soit un automorphisme. Alors si B est une base, $f(B)$ est une base. Par conséquent $\det(f) = \det_B\big(f(B)\big) \neq 0$ parce que $f(B)$ est une base (lemme 9.7(7)).
Réciproquement, supposons que $\det(f) \neq 0$. Alors si B est une base quelconque nous avons $\det_B\big(f(B)\big) \neq 0$, ce qui est uniquement possible lorsque $f(B)$ est une base. L'application f transforme donc toute base en une base et est alors un automorphisme d'espace vectoriel.

(3) Vu que le déterminant de l'identité est 1 et que f est inversible, $1 = \det(f\circ f^{-1}) = \det(f)\det(f^{-1})$.

□

3. Endomorphisme inversible, définition 7.17.

9.1.4 Déterminant d'une matrice

La proposition 7.32 nous assure que toute application linéaire possède une matrice associée (dans une base). Vu de la matrice, voici quelque conséquences de la proposition 9.10.

Proposition 9.11.
Quelques propriétés du déterminant vu de la matrice.

(1) *Si on permute deux lignes ou deux colonnes d'une matrice, alors le déterminant change de signe.*
(2) *Si on multiplie une ligne ou une colonne d'une matrice par un nombre λ, alors le déterminant est multiplié par λ.*
(3) *Si deux lignes ou deux colonnes sont proportionnelles, alors le déterminant est nul.*
(4) *Si on ajoute à une ligne une combinaison linéaire des autres lignes, alors le déterminant ne change pas (idem pour les colonnes).*

Le déterminant d'une matrice et d'une application linéaire est la définition 9.9, et les principales propriétés algébriques sont données dans la proposition 9.10.

En dimension deux, le déterminant de la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est le nombre

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb. \tag{9.39}$$

Ce nombre détermine entre autres le nombre de solutions que va avoir le système d'équations linéaires associé à la matrice.

Pour une matrice 3×3, nous avons le même concept, mais un peu plus compliqué ; nous avons la formule

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{9.40}$$

9.1.5 Déterminant de Vandermonde

Proposition 9.12 ([34])**.**
Le ***déterminant de Vandermonde*** *est le polynôme en n variables donné par*

$$V(T_1, \dots, T_n) = \det \begin{pmatrix} 1 & 1 & \dots & 1 \\ T_1 & T_2 & \dots & T_n \\ \vdots & \ddots & \ddots & \vdots \\ T_1^{n-1} & T_2^{n-1} & \dots & T_n^{n-1} \end{pmatrix} = \prod_{1 \leqslant i < j \leqslant n} (T_j - T_i). \tag{9.41}$$

Notez que l'inégalité du milieu est stricte (sinon d'ailleurs l'expression serait nulle).

Démonstration. Nous considérons le polynôme

$$f(X) = V(T_1, \dots, T_{n-1}, X) \in \big(\mathbb{K}[T_1, \dots, T_{n-1}]\big)[X]. \tag{9.42}$$

C'est un polynôme de degré au plus $n-1$ en X et il s'annule aux points $T_1, \dots, T_{n-1}$. Par conséquent il existe $\alpha \in \mathbb{K}[T_1, \dots, T_{n-1}]$ tel que

$$f = \alpha(X - T_{n-1}) \dots (X - T_1). \tag{9.43}$$

Nous trouvons α en écrivant $f(0)$. D'une part la formule (9.43) nous donne

$$f(0) = \alpha(-1)^{n-1} T_1 \dots T_{n-1}. \tag{9.44}$$

D'autre par la définition donne

$$f(0) = \det \begin{pmatrix} 1 & \cdots & 1 & 1 \\ T_1 & & T_{n-1} & 0 \\ \vdots & & \vdots & \vdots \\ T_1^{n-1} & \cdots & T_{n-1}^{n-1} & 0 \end{pmatrix} \tag{9.45a}$$

$$= (-1)^{n-1} \det \begin{pmatrix} T_1 & \ldots & T_{n-1} \\ \vdots & \ddots & \vdots \\ T_1^{n-1} & \ldots & T_{n-1}^{n-1} \end{pmatrix} \tag{9.45b}$$

$$= (-1)^{n-1} T_1 \ldots T_{n-1} \det \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ T_1^{n-1} & \cdots & T_{n-1}^{n-1} \end{pmatrix} \tag{9.45c}$$

$$= (-1)^{n-1} T_1 \ldots T_{n-1} V(T_1, \ldots, T_{n-1}) \tag{9.45d}$$

En égalisant avec (9.44), nous trouvons $\alpha = V(T_1, \ldots, T_{n-1})$, et donc

$$f = V(T_1, \ldots, T_{n-1}) \prod_{j \leqslant n-1} (X - T_j) \tag{9.46}$$

Enfin, une récurrence montre que

$$V(T_1, \ldots, T_n) = f(T_n) \tag{9.47a}$$

$$= V(T_1, \ldots, T_{n-1}) \prod_{j \leqslant n-1} (T_n - T_j) \tag{9.47b}$$

$$= \prod_{k \leqslant n} \prod_{j \leqslant k-1} (T_k - T_j) \tag{9.47c}$$

$$= \prod_{1 \leqslant j < k \leqslant n} (T_i - T_j). \tag{9.47d}$$

□

Exemple 9.13
Le déterminant de Vandermonde (proposition 9.12) est alterné, semi-symétrique et non symétrique. Le fait qu'il soit alterné est le fait qu'il soit un déterminant. Étant donné qu'il est alterné, il est semi-symétrique parce que sur A_n, nous avons $\epsilon = 1$. Étant donné qu'il est alterné, il change de signe sous l'action des éléments impairs de S_n et n'est donc pas symétrique. △

Proposition 9.14.
Un polynôme semi-symétrique $f \in \mathbb{K}[T_1, \ldots, T_n]$ se décompose de façon unique en

$$f = P + VQ \tag{9.48}$$

où P et Q sont deux polynômes symétriques.

Démonstration. Nous commençons par prouver l'unicité en montrant que si $f = PVQ$ avec P et Q symétrique, alors P et Q sont donnés par des formules explicites en termes de f.

Si σ_1 et σ_2 sont deux permutations impaires de $\{1, \ldots, n\}$, alors $\sigma_1 \cdot f = \sigma_2 \cdot f$ parce que l'élément $\sigma_2^{-1}\sigma_1$ est pair (proposition 8.39), de telle sorte que $\sigma_2^{-1}\sigma_1 \cdot f = f$. Nous posons donc $g = \tau \cdot f$ où τ est une permutation impaire quelconque – par exemple une transposition.

Vu que V est alternée et que τ est une transposition nous avons

$$g = \tau \cdot f = P - VQ. \tag{9.49}$$

Donc $f+g=2P$ et $f-g=2VQ$. Cela donne P et Q en terme de f et g, et donc l'unicité.

Attention : cela ne donne pas un moyen de prouver l'existence parce que rien ne prouve pour l'instant que $f-g$ peut effectivement être écrit sous la forme VQ, c'est à dire que $f-g$ soit divisible par V. C'est cela que nous allons nous atteler à démontrer maintenant.

Nous commençons par prouver que $f+g$ est symétrique et $f-g$ alterné. Si σ est une transposition,

$$\sigma\cdot(f+g)=\sigma\cdot f+\sigma\tau\cdot f=g+f \tag{9.50}$$

parce que $\sigma\tau$ est pair. De la même façon,

$$\sigma\cdot(f-g)=g-f=\epsilon(\sigma)(f-g). \tag{9.51}$$

Dans les deux cas nous concluons en utilisant le fait que toute permutation est un produit de transpositions (proposition 8.38) et que ϵ est un homomorphisme.

Soient maintenant deux entiers $h<k$ dans $\{1,\dots,n\}$ et l'anneau

$$\big(\mathbb{K}[T_1,\dots,\hat{T}_k,\dots,T_n]\big)[T_k]. \tag{9.52}$$

Cet anneau contient le polynôme T_k-T_h où T_k est la variable et T_h est un coefficient. Nous faisons la division euclidienne de $f-g$ par T_k-T_h parce que nous avons dans l'idée de faire arriver le déterminant de Vandermonde et donc le produit de toutes les différences T_k-T_h :

$$f-g=(T_k-T_h)q+r \tag{9.53}$$

où $\deg_{T_k} r<1$, c'est à dire que r ne dépends pas de T_k. Nous revoyons maintenant l'égalité (9.53) dans $\mathbb{K}[T_1,\dots,T_n]$ et nous y appliquons la transposition τ_{kh}. Nous savons que $\tau_{kh}(f-g)=-(f-g)$ et $\tau_{kh}(T_k-T_h)=-(T_k-T_h)$, et donc

$$-(f-g)=-(T_k-T_h)\tau_{kh}\cdot q+\tau_{kh}\cdot r \tag{9.54}$$

où $\tau_{kh}\cdot r$ ne dépend pas de T_h. Nous appliquons à (9.54) l'application

$$\begin{aligned} t\alpha\colon\mathbb{K}[T_1,\dots,T_n]&\to\mathbb{K}[T_1,\dots,\hat{T}_k,\dots,T_n]\\ \alpha(PT_1,\dots,\hat{T}_k,\dots,T_n)&=P(T_1,\dots,T_h,\dots,T_n).\end{aligned} \tag{9.55}$$

Cette application vérifie $\alpha\big(\tau_{kh}\cdot r\big)=\alpha(r)$ et nous avons

$$-\alpha(f-g)=\alpha(r). \tag{9.56}$$

Puis en appliquant α à la relation $f-g=(T_k-T_h)q+r$, nous trouvons

$$\alpha(f-g)=\alpha(r), \tag{9.57}$$

et par conséquent $\alpha(r)=0$. Ici nous utilisons l'hypothèse de caractéristique différente de deux. Dire que $\alpha(r)=0$, c'est dire que r est divisible par T_k-T_h, mais r étant de degré zéro en T_k, nous avons $r=0$. Par conséquent T_k-T_h divise $f-g$ pour tout $h<k$, et nous pouvons définir un polynôme Q par

$$f-g=2Q\prod_{h<k}\prod_{k\leqslant n}(T_k-T_h)=2Q(T_1,\dots,T_n)V(T_1,\dots,T_n), \tag{9.58}$$

où nous avons utilisé la formule du déterminant de Vandermonde de la proposition 9.12.

Étant donné que $f+g$ est un polynôme symétrique, nous allons aussi poser $f+g=2P$ avec P symétrique.

Montrons à présent que Q est un polynôme symétrique. Soit $\sigma\in S_n$; vu que nous savons déjà que $f-g$ est alternée, nous avons

$$\sigma\cdot(f-g)=\epsilon(\sigma)(f-g)=\epsilon(\sigma)2QV, \tag{9.59}$$

Mais en appliquant σ à l'équation (9.58),

$$\sigma \cdot (f - g) = 2(\sigma \cdot V)(T_1, \dots, , T_n)(\sigma \cdot Q)(T_1, \dots, T_n) \tag{9.60a}$$
$$= 2\epsilon(\sigma)V(T_1, \dots, T_n)(\sigma \cdot Q)(T_1, \dots, T_n). \tag{9.60b}$$

En égalisant avec (9.59) et en se souvenant que l'anneau $\mathbb{K}[T_1, \dots, T_n]$ était intègre (théorème 4.103), nous simplifions par $2\epsilon(\sigma)V$ pour obtenir

$$Q = \sigma \cdot Q, \tag{9.61}$$

c'est à dire que Q est symétrique.

Au final nous avons $f + q = 2P$ et $f - g = 2VQ$ avec P et Q symétriques. En faisant la somme,

$$f = P + VQ. \tag{9.62}$$

□

9.1.6 Déterminant de Gram

Si $x_1, \dots, x_r$ sont des vecteurs d'un espace vectoriel, alors le **déterminant de Gram** est le déterminant

$$G(x_1, \dots, x_r) = \det\big(\langle x_i, x_j \rangle\big). \tag{9.63}$$

Notons que la matrice est une matrice symétrique.

Proposition 9.15.
Si F est un sous-espace vectoriel de base $\{x_1, \dots, x_n\}$ et si x est un vecteur, alors le déterminant de Gram est un moyen de calculer la distance entre x et F par

$$d(x, F)^2 = \frac{G(x, x_1, \dots, x_n)}{G(x_1, \dots, x_n)}. \tag{9.64}$$

9.1.7 Déterminant de Cauchy

Soient des nombres a_i et b_i $(i = 1, \dots, n)$ tels que $a_i + b_j \neq 0$ pour tout couple (i, j). Le **déterminant de Cauchy** est

$$D_n = \det\left(\frac{1}{a_i + b_j}\right). \tag{9.65}$$

Proposition 9.16 ([95])**.**
Le déterminant de Cauchy est donné par la formule

$$D_n = \frac{\prod_{i<j}(a_j - a_i)\prod_{i<j}(b_j - b_i)}{\prod_{ij}(a_i + b_j)}. \tag{9.66}$$

9.1.8 Matrice de Sylvester

La définition est pompée de wikipédia. Soient P et Q deux polynômes non nuls, de degrés respectifs m et n :

$$P(x) = p_0 + p_1 x + \cdots + p_n x^n \tag{9.67a}$$
$$Q(x) = q_0 + q_1 x + \cdots + q_m x^m. \tag{9.67b}$$

La **matrice de Sylvester** associée à P et Q est la matrice carrée $m + n \times m + n$ définie ainsi :

(1) la première ligne est formée des coefficients de P, suivis de 0 :

$$\begin{pmatrix} p_n & p_{n-1} & \cdots & p_1 & p_0 & 0 & \cdots & 0 \end{pmatrix}; \tag{9.68}$$

(2) la seconde ligne s'obtient à partir de la première par permutation circulaire vers la droite ;
(3) les $(m-2)$ lignes suivantes s'obtiennent en répétant la même opération ;
(4) la ligne $(m+1)$ est formée des coefficients de Q, suivis de 0 :

$$\begin{pmatrix} q_m & q_{m-1} & \cdots & q_1 & q_0 & 0 & \cdots & 0 \end{pmatrix}; \tag{9.69}$$

(5) les $(m-1)$ lignes suivantes sont formées par des permutations circulaires.

Ainsi dans le cas $n=4$ et $m=3$, la matrice obtenue est

$$S_{p,q} = \begin{pmatrix} p_4 & p_3 & p_2 & p_1 & p_0 & 0 & 0 \\ 0 & p_4 & p_3 & p_2 & p_1 & p_0 & 0 \\ 0 & 0 & p_4 & p_3 & p_2 & p_1 & p_0 \\ q_3 & q_2 & q_1 & q_0 & 0 & 0 & 0 \\ 0 & q_3 & q_2 & q_1 & q_0 & 0 & 0 \\ 0 & 0 & q_3 & q_2 & q_1 & q_0 & 0 \\ 0 & 0 & 0 & q_3 & q_2 & q_1 & q_0 \end{pmatrix}. \tag{9.70}$$

Le déterminant de la matrice de Sylvester associée à P et Q est appelé le **résultant** de P et Q et noté $\operatorname{res}(P,Q)$.

Attention : si P est de degré n et Q de degré m, il y a m lignes pour P et n pour Q dans le déterminant du résultant (et non le contraire).

Lemme 9.17 ([96])**.**
Si P et Q sont deux polynômes de degrés n et m à coefficients dans l'anneau $\mathbb{A}$, alors pour tout $\lambda \in \mathbb{A}$,

$$\operatorname{res}(\lambda P, Q) = \lambda^m \operatorname{res}(P,Q) \tag{9.71a}$$
$$\operatorname{res}(P, \lambda Q) = \lambda^n \operatorname{res}(P,Q). \tag{9.71b}$$

Démonstration. Cela est simplement un comptage de nombre de lignes. Il y a m lignes contenant les coefficients de P ; donc prendre λP revient à multiplier m lignes dans un déterminant et donc le multiplier par λ^m. □

L'équation de Bézout (5.66) peut être traitée avec une matrice de Sylvester. Soient P et Q, deux polynômes donnés et à résoudre l'équation

$$xP + yQ = 0 \tag{9.72}$$

par rapport aux polynômes inconnus x et y dont les degrés sont $\deg(x) < \deg(Q)$ et $\deg(y) < \deg(P)$. Si nous notons $\tilde{x}$ et $\tilde{y}$ la liste des coefficients de x et y (dans l'ordre décroissant de degré), nous pouvons récrire l'équation (9.72) sous la forme

$$S_{PQ}^t \begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = 0. \tag{9.73}$$

Pour s'en convaincre, écrivons pour les polynômes de l'exemple (9.70) :

$$\begin{pmatrix} p_4 & 0 & 0 & q_3 & 0 & 0 & 0 \\ p_3 & p_4 & 0 & q_2 & q_3 & 0 & 0 \\ p_2 & p_3 & p_4 & q_1 & q_2 & q_3 & 0 \\ p_1 & p_2 & p_3 & q_0 & q_1 & q_2 & q_3 \\ p_0 & p_1 & p_2 & 0 & q_0 & q_1 & q_2 \\ 0 & p_0 & p_1 & 0 & 0 & q_0 & q_1 \\ 0 & 0 & p_0 & 0 & 0 & 0 & q_0 \end{pmatrix} \begin{pmatrix} x_2 \\ x_1 \\ x_0 \\ y_3 \\ y_2 \\ y_1 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_2p_4 + y_2q_3 \\ p_3x_2 + p_4x_1 + q_2y_3 + q_3y_2 \\ \\ \\ \vdots \end{pmatrix} \tag{9.74}$$

Nous voyons que sur la ligne numéro k (en partant du bas et en numérotant de à partir de zéro) nous avons les produits p_ix_j et q_iy_j avec $i+j=k$. La colonne de droite représente donc bien les coefficients du polynôme $xP+yQ$.

Proposition 9.18.
Le résultant de deux polynômes est non nul si et seulement si les deux polynômes sont premiers entre eux.

Un polynôme P a une racine double en a si et seulement si P et P' ont a comme racine commune, ce qui revient à dire que P et P' ne sont pas premiers entre eux.

Une application importante de ces résultats sera le théorème de Rothstein-Trager 14.90 sur l'intégration de fractions rationnelles.

Exemple 9.19
Si nous prenons $P = aX^2 + bX + c$ et $P' = 2aX + b$ alors la taille de la matrice de Sylvester sera $2 + 1 = 3$ et

$$S_{P,P'} = \begin{pmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{pmatrix}. \tag{9.75}$$

Le résultant est alors

$$\operatorname{res}(P, P') = -a(b^2 - 4ac). \tag{9.76}$$

Donc un polynôme du second degré a une racine double si et seulement si $b^2 - 4ac = 0$. Cela est un résultat connu depuis longtemps mais qui fait toujours plaisir à revoir. △

La matrice de Sylvester permet aussi de récrire l'équation de Bézout pour les polynômes ; voir le théorème 5.40 et la discussion qui s'ensuit.

Une proposition importante du résultant est qu'il peut s'exprimer à l'aide des racines des polynômes.

Proposition 9.20.
Si

$$P(X) = a_p \prod_{i=1}^{p} (X - \alpha_i) \tag{9.77a}$$

$$Q(X) = b_q \prod_{j=1}^{q} (X - \beta_i) \tag{9.77b}$$

alors nous avons les expressions suivantes pour le résultant :

$$\operatorname{res}(P, Q) = a_p^q b_q^p \prod_{i=1}^{p} \prod_{j=1}^{q} (\beta_j - \alpha_i) = b_q^p \prod_{j=1}^{q} P(\beta_j) = (-1)^{pq} a_p^q \prod_{i=1}^{p} Q(\alpha_i). \tag{9.78}$$

Démonstration. Si P et Q ne sont pas premiers entre eux, d'une part la proposition 9.18 nous dit que $\operatorname{res}(P, Q) = 0$ et d'autre part, P et Q ont un facteur irréductible en commun, ce qui signifie que nous devons avoir un des $X - \alpha_i$ égal à un des $X - \beta_j$. Autrement dit, nous avons $\alpha_i = \beta_j$ pour un couple (i, j). Par conséquent tous les membres de l'équation (9.78) sont nuls.

Nous supposons donc que P et Q sont premiers entre eux. Nous commençons par supposer que les polynômes P et Q sont unitaires, c'est à dire que $a_p = b_q = 1$. Nous considérons alors l'anneau

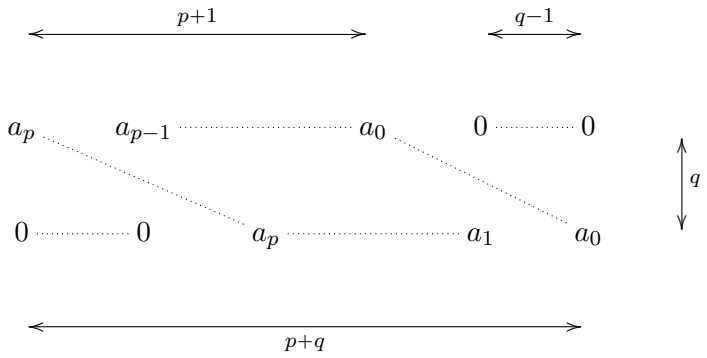
$$\mathbb{A} = \mathbb{Z}[\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q]. \tag{9.79}$$

Dans cet anneau, l'élément $\beta_j - \alpha_i$ est irréductible (tout comme $X - Y$ est irréductible dans $\mathbb{Z}[X, Y]$). Le résultant $R = \operatorname{res}(P, Q)$ est un élément de $\mathbb{A}$ parce que tous leurs coefficients peuvent être exprimés à l'aide des α_i et des β_j. Dans $\mathbb{A}$, l'élément $\beta_j - \alpha_i$ divise R. En effet lorsque $\beta_j = \alpha_i$, le déterminant définissant le résultant est nul, ce qui signifie que $\beta_j - \alpha_i$ est un facteur irréductible de R.

Par conséquent il existe un polynôme $T \in \mathbb{A}$ tel que

$$R = \lambda(\alpha_1, \dots, \beta_q) \prod_{i=1}^{p} \prod_{j=1}^{r} (\beta_j - \alpha_i). \tag{9.80}$$

Comptons les degrés. Pour donner une idée de ce calcul de degré, voici comment se présente, au niveau des dimensions, le déterminant :



(9.81)

si les a_i sont les coefficients de P. Mais chacun des a_i est de degré 1 en les α_i, donc le déterminant dans son ensemble est de degré q en les α_i, parce que R contient q lignes telles que (9.81). Le même raisonnement montre que R est de degré p en les β_j. Par ailleurs le polynôme $\prod_{i=1}^{p}\prod_{j=1}^{r}(\beta_j-\alpha_i)$ est de degré p en les β_j et q en les α_i. Nous en déduisons que T doit être un polynôme ne dépendant pas de α_i ou de β_j.

Nous pouvons donc calculer la valeur de T en choisissant un cas particulier. Avec $P(X)=X^p$ et $Q(X)=X^q+1$, il est vite vu que $R(P,Q)=1$ et donc que $T=1$.

Si les polynômes P et Q ne sont pas unitaires, le lemme 9.17 nous permet de conclure.

□

9.1.9 Théorème de Kronecker

Nous considérons K_n l'ensemble des polynômes de $\mathbb{Z}[X]$

(1) unitaires de degré n,

(2) dont les racines dans $\mathbb{C}$ sont de modules plus petits ou égaux à 1,

(3) et qui ne sont pas divisés par X.

Un tel polynôme s'écrit sous la forme

$$P=X^n+\sum_{k=0}^{n-1}a_kX^k. \tag{9.82}$$

Théorème 9.21 (Kronecker[2])**.**
Les racines des éléments de K_n sont des racines de l'unité.

Démonstration. Vu que $\mathbb{C}$ est algébriquement clos nous pouvons considérer les racines $\alpha_1,\dots,\alpha_n$ de P dans $\mathbb{C}$. Nous les considérons avec leurs multiplicités.

Soit $R=X^n+\sum_{k=0}^{n-1}b_kX^k$ un élément de K_n dont nous notons $\beta_1,\dots,\beta_n$ les racines dans $\mathbb{C}$. Les relations coefficients-racines stipulent que

$$b_k=\sum_{1\leqslant i_1<\dots<i_{n-k}\leqslant n}\prod_{j=1}^{n-k}\beta_{i_j}. \tag{9.83}$$

En prenant le module et en se souvenant que $|\beta_l|\leqslant 1$ pour tout l, nous trouvons que

$$|b_k|\leqslant\binom{n}{n-k}. \tag{9.84}$$

Mais comme $b_k\in\mathbb{Z}$, nous avons

$$b_k\in\Big\{-\binom{n}{n-k},-\binom{n}{n-k}+1,\dots,0,\cdots,\binom{n}{n-k}\Big\} \tag{9.85}$$

qui est de cardinal $\binom{n}{n-k}+1$. Nous avons donc

$$\text{Card}(K_n) \leqslant \prod_{k=0}^{n-1} \left(1+\binom{n}{n-k}\right) < \infty. \tag{9.86}$$

La conclusion jusqu'ici est que K_n est un ensemble fini.

Pour chaque $k \in \mathbb{N}^*$ nous considérons les polynômes

$$P_k = \prod_{i=1}^{n}(X-\alpha_i^k) \tag{9.87a}$$

$$Q_k = X^k - Y \in \mathbb{Z}[X,Y], \tag{9.87b}$$

et puis nous considérons le résultant $R_k = \text{res}_X(P,Q_k) \in \mathbb{Z}[Y]$:

$$R_k = \text{res}_X(P,Q_k) = \begin{pmatrix} 1 & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & & & \\ 0 & \cdots & 0 & 1 & a_{n-1} & \cdots & a_0 & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & a_{n-1} & \cdots & a_0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 & a_{n-1} & \cdots & a_0 \\ & & & & & & & & \\ 1 & 0 & \ldots & 0 & -Y & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & -Y & 0 & \ldots & 0 \\ & & \ddots & & & \ddots & \ddots & & \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -Y & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & -Y \end{pmatrix} \tag{9.88}$$

Cela est un polynôme en Y dont le terme de plus haut degré est $(-1)^nY^n$. Les petites formules de la proposition 9.20 nous permettent d'exprimer $R_k(Y)$ en termes des racines de P :

$$R_k(Y) = \prod_{i=1}^{n} Q_k(\alpha_i) = \prod_{i=1}^{n}(\alpha_i^k - Y) = (-1)^n \prod_{i=1}^{n}(Y-\alpha_i^k) = (-1)^n P_k(Y). \tag{9.89}$$

Vu que $P \in K_n$ nous savons que les α_i ne sont pas tous nuls ; donc $P_k \in K_n$. Cependant nous avons vu que K_n est un ensemble fini ; donc parmi les P_k, il y a des doublons (et pas un peu)[4] . Nous regardons même l'ensemble des P_{2^n} dans lequel nous pouvons en trouver deux les mêmes. Soit $l > k$ tels que $P_{2^k} = P_{2^l}$. Si α est racine de P_{2^k}, alors il est de la forme $\alpha = \beta^{2^k}$ pour une certaine racines β de P. Par conséquent

$$\alpha^{2^l/2^k} = \alpha^{2^{l-k}} \tag{9.90}$$

est racine de P_{2^l}. Notons que dans cette expression il n'y a pas de problèmes de définition d'exposant fractionnaire dans $\mathbb{C}$ parce que $l > k$. Vu que (9.90) est racine de P_{2^l}, il est aussi racine de P_{2^k}. Donc

$$\left(\alpha^{2^{l-k}}\right)^{2^{l-k}} = \alpha^{2^{2(l-k)}} \tag{9.91}$$

est racine de P_{2^l} et donc de P_{2^k}. Au final nous savons que tous les nombres de la forme $\alpha^{2^{n(l-k)}}$ sont racines de P_{2^k}. Mais comme P_{2^k} a un nombre fini de racines, nous pouvons en trouver deux égales. Si nous avons

$$\alpha^{2^{n(l-k)}} = \alpha^{2^{m(l-k)}} \tag{9.92}$$

pour certains entiers $m > n$, alors

$$\alpha^{2^{n(l-k)}-2^{m(l-k)}} = 1, \tag{9.93}$$

ce qui prouver que α est une racine de l'unité. Nous avons donc prouvé que toutes les racines de P_{2^k} sont des racines de l'unité et donc que les racines de P sont racines de l'unité. □

4. Ici dans [2], il déduit qu'on a un k tel que $P_k = P_1 = P$. Mois je vois pourquoi on a un k et un l tels que $P_k = P_l$, mais pourquoi on peut en trouver un spécialement égal au premier ? Une réponse à cette question permettrait de solidement réduire la lourdeur de la suite de la preuve.

9.2 Produit scalaire

Définition 9.22.
*Une **forme bilinéaire** sur un espace vectoriel E est une application $b\colon E\times E\to\mathbb{K}$ telle que*

(1) $b(u,v)=b(v,u)$,

(2) $b(u+v,w)=b(u,w)+b(v,w)$,

(3) $b(\lambda u,v)=\lambda b(u,v)$

pour tout $u,v,w\in E$ et $\lambda\in\mathbb{K}$ où $\mathbb{K}$ est un corps commutatif.

Définition 9.23.
Si g est une application bilinéaire sur un espace vectoriel E nous disons qu'elle est

(1) ***définie positive*** *si $g(u,u)\geqslant 0$ pour tout $u\in E$ et $g(u,u)=0$ si et seulement si $u=0$.*

(2) ***semi-définie positive*** *si $g(u,u)\geqslant 0$ pour tout $u\in E$. Nous dirons aussi parfois qu'elle est simplement «positive».*

Cela est évidemment à lier à la définition 9.167 et la proposition 9.247 : une application bilinéaires est définie positive si et seulement si sa matrice symétrique associée l'est.

Définition 9.24.
*Un **produit scalaire** sur un espace vectoriel E est une forme bilinéaire*[5] *symétrique strictement définie positive*[6].

Étant donné que l'inégalité de Cauchy-Schwarz sera surtout utilisée dans le cas où un produit scalaire est bel et bien donné, nous l'énonçons et le démontrons avec des notations adaptée à l'usage. Le produit scalaire sera noté $X\cdot Y$ pour $b(X,Y)$ si b est la forme bilinéaire.

Théorème 9.25 (Inégalité de Cauchy-Schwarz)**.**
Si X et Y sont des vecteurs, alors

$$|X\cdot Y|\leqslant\|X\|\|Y\|. \tag{9.94}$$

Nous avons une égalité si et seulement si X et Y sont multiples l'un de l'autre.

Démonstration. Étant donné que les deux membres de l'inéquation sont positifs, nous allons travailler en passant au carré afin d'éviter les racines carrés dans le second membre.

Nous considérons le polynôme

$$P(t)=\|X+tY\|^2=(X+tY)\cdot(X+tY)=X\cdot X+tX\cdot Y+tY\cdot X+t^2Y\cdot Y. \tag{9.95}$$

En ordonnant les termes selon les puissance de t,

$$P(t)=\|Y\|^2t^2+2(X\cdot Y)t+\|X\|^2. \tag{9.96}$$

Cela est un polynôme du second degré en t. Par conséquent le discriminant[7] doit être négatif. Nous avons donc

$$\Delta=4(X\cdot Y)^2-4\|X\|^2\|Y\|^2\leqslant 0, \tag{9.97}$$

ce qui donne immédiatement

$$(X\cdot Y)^2\leqslant\|X\|^2\|Y^2\|. \tag{9.98}$$

En ce qui concerne le cas d'égalité, si nous avons $X\cdot Y=\|X\|\|Y\|$, alors le discriminant Δ ci-dessus est nul et le polynôme P admet une racine double t_0. Pour cette valeur nous avons

$$P(t_0)=|X+t_0Y|=0, \tag{9.99}$$

ce qui implique $X+t_0Y=0$ et donc que X et Y sont liés. □

5. Définition 9.22.
6. Définition 9.23.
7. Le fameux b^2-4ac.

Vu que nous allons voir un pâté d'espaces avec des produits scalaires, nous leur donnons un nom.

Définition 9.26.
Un espace vectoriel ***euclidien*** *est un espace vectoriel de dimension finie muni d'un produit scalaire (définition 9.24).*

Définition 9.27 (Isométrie, thème 61)**.**
Une ***isométrie*** *d'une forme bilinéaire b sur l'espace vectoriel E est une application bijective $f\colon E \to E$ telle que $b\big(f(u), f(v)\big) = b(u, v)$ pour tout $u, v \in E$.*

En particulier une isométrie d'un espace euclidien est une application bijective qui préserve le produit scalaire.

Lemme 9.28.
Une isométrie d'un espace euclidien fixe l'origine.

Démonstration. Soit une isométrie f d'un espace euclidien : $f(x) \cdot f(y) = x \cdot y$ pour tout $x, y \in E$. En particulier pour $x = 0$ nous avons

$$f(0) \cdot f(y) = 0 \tag{9.100}$$

pour tout y. Vu que f est une bijection, nous avons $f(0) \cdot x = 0$ pour tout x. Comme le produit scalaire est non dégénéré cela implique que $f(0) = 0$. □

Proposition 9.29.
Si $x, y \mapsto x \cdot y$ est un produit scalaire sur un espace vectoriel E, alors $N(x) = \sqrt{x \cdot x}$ est une norme vérifiant l'identité du parallélogramme :

$$\|x - y\|^2 + \|x + y\|^2 = 2\|x\|^2 + 2\|y\|^2. \tag{9.101}$$

Démonstration. Prouvons l'inégalité triangulaire. Si $x, y \in E$ nous avons

$$\|x + y\| = \sqrt{\|x\|^2 + \|y\|^2 + 2x \cdot y}. \tag{9.102}$$

Par l'inégalité de Cauchy-Schwarz, théorème 9.25 nous avons aussi

$$\|x\|^2 + \|y\|^2 + 2x \cdot y \leqslant \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = \big(\|x\| + \|y\|\big)^2, \tag{9.103}$$

donc

$$\|x + y\| \leqslant \sqrt{\big(\|x\| + \|y\|\big)^2} = \|x\| + \|y\|. \tag{9.104}$$

La seconde assertion est seulement un calcul :

$$\begin{aligned}\|x - y\|^2 + \|x + y\|^2 &= (x - y) \cdot (x - y) + (x + y) \cdot (x + y) \\ &= x \cdot x - x \cdot y - y \cdot x + y \cdot y \\ &\quad + x \cdot x + x \cdot y + y \cdot x + y \cdot y \\ &= 2x \cdot x + 2y \cdot y \\ &= 2\|x\|^2 + 2\|y\|^2.\end{aligned} \tag{9.105}$$

□

Le produit scalaire permet de donner une norme via la formule suivante :

$$\|x\|^2 = x \cdot x. \tag{9.106}$$

Lemme 9.30 ([2])**.**
Soit V un espace vectoriel muni d'un produit scalaire et de la norme associée. Si $x, y \in V$ satisfont à $\|x + y\| = \|x\| + \|y\|$, alors il existe $\lambda \geqslant 0$ tel que $x = \lambda y$.

Démonstration. Quitte à raisonner avec $x/\|x\|$ et $y/\|y\|$, nous supposons que $\|x\| = \|y\| = 1$. Dans ce cas l'hypothèse signifie que $\|x+y\|^2 = 4$. D'autre part en écrivant la norme en terme de produit scalaire,

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle, \tag{9.107}$$

ce qui nous mène à affirmer que $\langle x, y\rangle = 1 = \|x\|\|y\|$. Nous sommes donc dans le cas d'égalité de l'inégalité de Cauchy-Schwarz [8], ce qui nous donne un λ tel que $x = \lambda y$. Étant donné que $\|x\| = \|y\| = 1$ nous avons obligatoirement $\lambda = \pm 1$, mais si $\lambda = -1$ alors $\langle x, y\rangle = -1$, ce qui est le contraire de ce qu'on a prétendu plus haut. Par soucis de cohérence, nous allons donc croire que $\lambda = 1$. □

Proposition 9.31.
si $v_1, \cdots, v_k$ sont des vecteurs non nuls, orthogonaux deux à deux, alors ces vecteurs forment une famille libre.

9.2.1 Projection et angles

Proposition 9.32 (Propriétés du produit scalaire)**.**
Si X et Y sont des vecteurs de $\mathbb{R}^3$, alors

Symétrie $X \cdot Y = Y \cdot X$;

Linéarité $(\lambda X + \mu X') \cdot Y = \lambda(X \cdot Y) + \mu(X' \cdot Y)$ *pour tout* λ *et* μ *dans* $\mathbb{R}$;

Défini positif $X \cdot X \geqslant 0$ *et* $X \cdot X = 0$ *si et seulement si* $X = 0$.

Note : lorsque nous écrivons $X = 0$, nous voulons voulons dire $X = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Définition 9.33.
La ***norme*** *du vecteur X, notée $\|X\|$, est définie par*

$$\|X\| = \sqrt{X \cdot X} = \sqrt{x^2 + y^2 + z^2} \tag{9.108}$$

si $X = (x, y, z)$. Cette norme sera parfois nommée «norme euclidienne».

Cette définition est motivée par le théorème de Pythagore. Le nombre $X \cdot X$ est bien la longueur de la «flèche» X. Plus intrigante est la définition suivante :

Définition 9.34.
Deux vecteurs X et Y sont ***orthogonaux*** *si $X \cdot Y = 0$.*

Cette définition de l'orthogonalité est motivée par la proposition suivante.

Proposition 9.35.
Si nous écrivons $\texttt{proj}_Y$ l'opération de projection sur la droite qui sous-tend Y, alors nous avons

$$\|\texttt{proj}_Y X\| = \frac{X \cdot Y}{\|Y\|}. \tag{9.109}$$

Démonstration. Les vecteurs X et Y sont des flèches dans l'espace. Nous pouvons choisir un système d'axe orthogonal tel que les coordonnées de X et Y soient

$$X = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \quad Y = \begin{pmatrix} l \\ 0 \\ 0 \end{pmatrix} \tag{9.110}$$

où l est la longueur du vecteur Y. Pour ce faire, il suffit de mettre le premier axe le long de Y, le second dans le plan qui contient X et Y, et enfin le troisième axe dans le plan perpendiculaire aux deux premiers.

8. Théorème 9.25.

Un simple calcul montre que $X \cdot Y = xl + y \cdot 0 + 0 \cdot 0 = xl$. Par ailleurs, nous avons $\| \texttt{proj}_Y X \| = x$. Par conséquent,

$$\| \texttt{proj}_Y X \| = \frac{X \cdot Y}{l} = \frac{X \cdot Y}{\|Y\|}. \tag{9.111}$$

□

Corollaire 9.36.
Si la norme de Y est 1, alors le nombre $X \cdot Y$ est la longueur de la projection de X sur Y.

Démonstration. Poser $\|Y\| = 1$ dans la proposition 9.35. □

Remarque 9.37.
Outre l'orthogonalité, le produit scalaire permet de savoir l'angle entre deux vecteurs à travers la définition 16.47. D'autres interprétations géométriques du déterminant sont listées dans le thème 52.

Nous sommes maintenant en mesure de déterminer, pour deux vecteurs quelconques u et v, la projection orthogonale de u sur v. Ce sera le vecteur $\bar{u}$ parallèle à v tel que $u - \bar{u}$ est orthogonal à v. Nous avons donc

$$\bar{u} = \lambda v \tag{9.112}$$

et

$$(u - \lambda v) \cdot v = 0. \tag{9.113}$$

La seconde équation donne $u \cdot v - \lambda v \cdot v = 0$, ce qui fournit λ en fonction de u et v :

$$\lambda = \frac{u \cdot v}{\|v\|^2}. \tag{9.114}$$

Nous avons par conséquent

$$\bar{u} = \frac{u \cdot v}{\|v\|^2} v. \tag{9.115}$$

Armés de cette interprétation graphique du produit scalaire, nous comprenons pourquoi nous disons que deux vecteurs sont orthogonaux lorsque leur produit scalaire est nul.

Nous pouvons maintenant savoir quel est le coefficient directeur d'une droite orthogonale à une droite donnée. En effet, supposons que la première droite soit parallèle au vecteur X et la seconde au vecteur Y. Les droites seront perpendiculaires si $X \cdot Y = 0$, c'est à dire si

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 0. \tag{9.116}$$

Cette équation se développe en

$$x_1 y_1 = -x_2 y_2. \tag{9.117}$$

Le coefficient directeur de la première droite est $\frac{x_2}{x_1}$. Isolons cette quantité dans l'équation (9.117) :

$$\frac{x_2}{x_1} = -\frac{y_1}{y_2}. \tag{9.118}$$

Donc le coefficient directeur de la première est l'inverse et l'opposé du coefficient directeur de la seconde.

Exemple 9.38
Soit la droite $d \equiv y = 2x + 3$. Le coefficient directeur de cette droite est 2. Donc le coefficient directeur d'une droite perpendiculaires doit être $-\frac{1}{2}$. △

Preuve alternative. La preuve peut également être donnée en ne faisant pas référence au produit scalaire. Il suffit d'écrire toutes les quantités en termes des coordonnées de X et Y. Si nous posons

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \tag{9.119}$$

l'inégalité à prouver devient

$$(x_1y_1 + x_2y_2 + x_3y_3)^2 \leqslant (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2). \tag{9.120}$$

Nous considérons la fonction

$$\varphi(t) = (x_1 + ty_1)^2 + (x_2 + ty_2)^2 + (x_3 + ty_3)^2 \tag{9.121}$$

En tant que norme, cette fonction est évidement positive pour tout t. En regroupant les termes de chaque puissance de t, nous avons

$$\varphi(t) = (y_1^2 + y_2^2 + y_3^2)t^2 + 2(x_1y_1 + x_2y_2 + x_3y_3)t + (x_1^2 + x_2^2 + x_3^2). \tag{9.122}$$

Cela est un polynôme du second degré en t. Par conséquent le discriminant doit être négatif. Nous avons donc

$$4(x_1y_1 + x_2y_2 + x_3y_3)^2 - (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) \leqslant 0. \tag{9.123}$$

La thèse en découle aussitôt. □

Proposition 9.39.
La norme euclidienne a les propriétés suivantes :

(1) Pour tout vecteur X et réel λ, $\|\lambda X\| = |\lambda|\|X\|$. Attention à ne pas oublier la valeur absolue !
(2) Pour tout vecteurs X et Y, $\|X + Y\| \leqslant \|X\| + \|Y\|$.

Démonstration. Nous ne prouvons pas le premier point. Pour le second, nous avons les inégalités suivantes :

$$\begin{aligned} \|X + Y\|^2 &= \|X\|^2 + \|Y\|^2 + 2X \cdot Y && \text{(9.124a)} \\ &\leqslant \|X\|^2 + \|Y\|^2 + 2|X \cdot Y| && \text{(9.124b)} \\ &\leqslant \|X\|^2 + \|Y\|^2 + 2\|X\|\|Y\| && \text{(9.124c)} \\ &= \big(\|X\| + \|Y\|\big)^2 && \text{(9.124d)} \end{aligned}$$

Nous avons utilisé d'abord la majoration $|x| \geqslant x$ qui est évident pour tout nombre x ; et ensuite l'inégalité de Cauchy-Schwarz 9.25. □

9.2.2 Procédé de Gram-Schmidt

Proposition 9.40 (Procédé de Gram-Schmidt)**.**
Un espace euclidien possède une base orthonormée.

Démonstration. Soit E un espace euclidien et $\{v_1, \dots, v_n\}$, une base quelconque de E. Nous posons d'abord

$$f_1 = v_1, \quad e_1 = \frac{f_1}{\|f_1\|}. \tag{9.125}$$

Ensuite

$$f_2 = v_2 - \langle v_2, e_1\rangle e_1, \quad e_2 = \frac{f_2}{\|f_2\|}. \tag{9.126}$$

Notons que $\{e_1, e_2\}$ est une base de $\mathrm{Span}\{v_1, v_2\}$. De plus elle est orthogonale :

$$\langle e_1, f_2\rangle = \langle e_1, v_2\rangle - \langle v_2, e_1\rangle \underbrace{\langle e_1, e_1\rangle}_{=1} = 0. \tag{9.127}$$

Le fait que $\|e_1\| = \|e_2\| = 1$ est par construction. Nous avons donc donné une base orthonormée de $\text{Span}\{v_1, v_2\}$.

Nous continuons par récurrence en posant

$$f_k = v_k - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle e_i, \quad e_k = \frac{f_k}{\|f_k\|}. \tag{9.128}$$

Pour tout $j < k$ nous avons

$$\langle e_j, f_k \rangle = \langle e_j, v_k \rangle - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle \underbrace{\langle e_i, e_j \rangle}_{=\delta_{ij}} = 0 \tag{9.129}$$

□

Cet algorithme de Gram-Schmidt nous donne non seulement l'existence de bases orthonormée pour tout espace euclidien, mais aussi le moyen d'en construire à partir de n'importe quelle base.

9.3 Hermitien, orthogonal, adjoint

Définition 9.41.
*Si E est un espace vectoriel sur $\mathbb{C}$, nous disons qu'une application $\langle ., . \rangle : E \times E \to \mathbb{C}$ est un **produit scalaire hermitien** si pour tout $u, v \in E$ nous avons*

(1) $\langle u, v \rangle = \overline{\langle v, u \rangle}$

(2) $\lambda \langle u, v \rangle = \langle \lambda u, v \rangle = \langle u, \bar{\lambda} v \rangle$

(3) $\langle u, u \rangle \in \mathbb{R}^+$ *et* $\langle u, u \rangle = 0$ *si et seulement si* $u = 0$.

Un espace vectoriel sur $\mathbb{C}$ muni d'un produit scalaire hermitien est un espace vectoriel hermitien.

Exemple 9.42
L'ensemble $E = \mathbb{C}^n$ vu comme espace vectoriel de dimension n sur $\mathbb{C}$ est muni d'une forme sesquilinéaire

$$\langle x, y \rangle = \sum_{k=1}^{n} x_k \bar{y}_k \tag{9.130}$$

pour tout $x, y \in \mathbb{C}^n$. Cela est un espace vectoriel hermitien. △

Définition 9.43.
*Soit E un espace vectoriel hermitien et $A \in \text{End}(E)$. L'opérateur **adjoint** de A est l'opérateur noté A^* définit par*

$$\langle Av, w \rangle = \langle v, A^* w \rangle \tag{9.131}$$

pour tout $v, w \in \mathbb{C}^n$.

La prescription (9.131) définit bien l'opérateur A^* de façon unique. En effet si $\{e_i\}$ est une base orthonormée alors

$$\langle e_i, A^* e_j \rangle = (A^* e_j)_i = \langle Ae_i, e_j \rangle = (Ae_i)_j. \tag{9.132}$$

Autrement dit, $(A^* e_j)_i = (Ae_i)_j$.

Proposition 9.44.
En ce qui concerne le déterminant,

$$\det(A^*) = \det(A)^* \tag{9.133}$$

où l'étoile à droite dénote la conjugaison complexe dans $\mathbb{C}$.

Démonstration. Écrivons la définition (9.17) avec l'expression explicite (9.17) du déterminant. Le tout avec la base canonique :

$$\det(A) = \det_{(e_1,\dots,e_n)}(Ae_1,\dots,Ae_n) = \sum_{\sigma\in S_n}\epsilon(\sigma)\prod_{i=1}^{n} e^*_{\sigma(i)}(Ae_i). \tag{9.134}$$

Mais nous pouvons développer :

$$e^*_{\sigma(i)}(Ae_i) = \langle e_{\sigma(i)}, Ae_i\rangle = \langle A^*e_{\sigma(i)}, e_i\rangle = \langle e_i, A^*e_{\sigma(i)}\rangle^* = e^*_i(A^*e_{\sigma(i)})^*. \tag{9.135}$$

Notez que dans la dernière expression, les trois * ont trois significations différentes. Par conséquent,

$$\det(A) = \sum_{\sigma\in S_n}\epsilon(\sigma)\prod_{i=1}^{n} e^*_i(A^*e_{\sigma(i)})^*. \tag{9.136}$$

Mais $e^*_i(A^*e_{\sigma(i)}) = e^*_{\sigma(j)}(A^*e_j)$ pour $j=\sigma(i)$, donc le produit ne change pas si on déplace le σ :

$$\det(A) = \sum_{\sigma\in S_n}\epsilon(\sigma)\prod_{i=1}^{n} e^*_{\sigma(i)}(A^*e_i)^* = \det(A^*)^*. \tag{9.137}$$

□

Définition 9.45.
*Un opérateur A est **hermitien** si $A^* = A$. On dit aussi **autoadjoint**.*

Le mot «hermitien» est réservé aux opérateurs sur des espaces hermitiens, c'est à dire des espaces vectoriels sur $\mathbb{C}$. Le mot «autoadjoint» par contre est plutôt utilisé dans le cadre d'opérateurs sur les espaces réels[9]. En conséquence de quoi, ces deux mots sont synonymes, mais il est préférable d'utiliser «hermitien» lorsque l'espace vectoriel est sur $\mathbb{C}$ et «autoadjoint» lorsqu'il est sur $\mathbb{R}$.

Remarque 9.46.
Le fait d'être hermitien n'implique en rien le fait d'être inversible.

9.3.1 Opérateur orthogonal, matrice orthogonale

Définition 9.47.
*Un opérateur est **orthogonal** lorsque $A^* = A^{-1}$ où A^* est l'adjoint de A définit en 9.43.*

Définition 9.48.
*Une matrice U est **orthogonale** si $U^t = U^{-1}$. Le **groupe orthogonal** noté $\mathrm{O}(n)$ est l'ensemble des matrices orthogonales $n\times n$.*

Lemme 9.49.
Soit un opérateur $A\colon \mathbb{R}^n \to \mathbb{R}^n$ muni du produit scalaire usuel. Il est orthogonal si et seulement si sa matrice dans la base canonique est orthogonale[10].

Démonstration. Soit la base canonique $\{e_i\}_{i=1,\dots,n}$ de $\mathbb{R}^n$. Nous avons

$$\langle AA^*e_i, e_j\rangle = \langle e_i, e_j\rangle = \delta_{ij}, \tag{9.138}$$

donc $\big((AA^*)e_j\big)_j = \delta_{ij}$, ou encore $(AA^*)_{ij} = \delta_{ij}$, ce qui signifie que la matrice AA^* est l'identité. □

Proposition 9.50 (Thème 61)**.**
À propos de matrices orthogonales.

9. Voir la définition 9.260.
10. Définition 9.48.

(1) L'ensemble des matrice réelles orthogonales forme un groupe noté $\mathrm{O}(n,\mathbb{R})$.

(2) Si A est une matrice orthogonale, alors $\det(A)=\pm 1$.

(3) Le groupe $\mathrm{O}(n)$ *est le groupe des isométries de* $\mathbb{R}^n$.

Démonstration. Si A et B sont orthogonales, alors

$$(AB)(AB)^t=ABB^tA^t=A\mathbb{1}A^t=\mathbb{1}. \tag{9.139}$$

Vu que $\mathbb{1}$ est orthogonale, nous avons bien un groupe.

En ce qui concerne le déterminant, $AA^t=\mathbb{1}$ donne $\det(A)\det(A^t)=1$, mais la proposition 9.44 dit que $\det(A)=\det(A^t)$, donc $\det(A)^2=1$. D'où le fait que $\det(A)=\pm 1$.

D'autre part si A est une isométrie de $\mathbb{R}^n$ alors pour tout $x,y\in\mathbb{R}^n$ nous avons $\langle Ax,Ay\rangle=\langle x,y\rangle$. En particulier,

$$\langle A^tAx,y\rangle=\langle x,y\rangle \tag{9.140}$$

pour tout $x,y\in\mathbb{R}^n$. En prenant $y=e_i$ nous trouvons

$$(A^tAx)_i=x_i, \tag{9.141}$$

ce qui signifie que pour tout x, $A^tAx=x$, ou encore que A^tA est l'identité.

Réciproquement si A^tA est l'identité nous avons

$$\langle a,y\rangle=\langle A^tAx,y\rangle=\langle Ax,Ay\rangle, \tag{9.142}$$

ce qui prouve que A est une isométrie. □

En ce qui concerne les valeurs propres des matrices de $\mathrm{O}(n)$ ainsi que leurs formes canoniques (avec des fonctions trigonométriques) pour O(3) et SO(3), ce sera pour la proposition 16.126 et ce qui s'ensuit.

Définition 9.51.
Le sous-groupe des matrices orthogonales de déterminant 1 *est le groupe* ***spécial orthogonal*** *noté* $\mathrm{SO}(n)$.

9.4 Directions conservées

9.4.1 Généralités

Nous savons qu'une application *linéaire* $A\colon\mathbb{R}^3\to\mathbb{R}^3$ est complètement définie par la donnée de son action sur les trois vecteurs de base, c'est à dire par la donnée de

$$Ae_1,\quad Ae_2\quad\text{et}\quad Ae_3. \tag{9.143}$$

Nous allons former la matrice de A en mettant simplement les vecteurs Ae_1, Ae_2 et Ae_3 en colonne. Donc la matrice

$$A=\begin{pmatrix}3&0&0\\0&1&0\\0&1&0\end{pmatrix} \tag{9.144}$$

signifie que l'application linéaire A envoie le vecteur e_1 sur $\begin{pmatrix}3\\0\\0\end{pmatrix}$, le vecteur e_2 sur $\begin{pmatrix}0\\0\\1\end{pmatrix}$ et le vecteur e_3 sur $\begin{pmatrix}0\\1\\0\end{pmatrix}$. Pour savoir comment A agit sur n'importe quel vecteur, on applique la règle de produit vecteur×matrice :

$$\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x+2y+3z\\4x+5y+6z\\7x+8y+9z\end{pmatrix}. \tag{9.145}$$

Une chose intéressante est de savoir quelles sont les directions invariantes de la transformation linéaire. Par exemple, on peut lire sur la matrice (9.144) que la direction $\begin{pmatrix}1\\0\\0\end{pmatrix}$ est invariante : elle est simplement multipliée par 3. Dans cette direction, la transformation est juste une dilatation. Afin de savoir si v est un vecteur d'une direction conservée, il faut voir s'il existe un nombre λ tel que $Av = \lambda v$, c'est à dire voir si v est simplement dilaté.

L'équation $Av = \lambda v$ se récrit $(A - \lambda\mathbb{1})v = 0$, c'est à dire qu'il faut résoudre l'équation

$$(A - \lambda\mathbb{1})\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}. \tag{9.146}$$

Nous savons qu'une telle équation ne peut avoir de solutions que si $\det(A - \lambda\mathbb{1}) = 0$. La première étape est donc de trouver les λ qui vérifient cette condition.

9.4.2 Comment trouver la matrice d'une symétrie donnée ?

9.4.2.1 Symétrie par rapport à un plan

Comment trouver par exemple la matrice A qui donne la symétrie autour du plan $z = 0$? La définition d'une telle symétrie est que les vecteurs du plan $z = 0$ ne bougent pas, tandis que les vecteurs perpendiculaires changent de signe. Ces informations vont permettre de trouver comment A agit sur une base de $\mathbb{R}^3$. En effet :

(1) Le vecteur $\begin{pmatrix}1\\0\\0\end{pmatrix}$ est dans le plan $z = 0$, donc il ne bouge pas,

(2) le vecteur $\begin{pmatrix}0\\1\\0\end{pmatrix}$ est également dans le plan, donc il ne bouge pas non plus,

(3) et le vecteur $\begin{pmatrix}0\\0\\1\end{pmatrix}$ est perpendiculaire au plan $z = 0$, donc il va changer de signe.

Cela nous donne directement les valeurs de A sur la base canonique et nous permet d'écrire

$$A = \begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{pmatrix}. \tag{9.147}$$

Pour écrire cela, nous avons juste mit en colonne les images des vecteurs de base. Les deux premiers n'ont pas changé et le troisième a changé.

Et si maintenant on donne un plan moins facile que $z = 0$? Le principe reste le même : il faudra trouver deux vecteurs qui sont dans le plan (et dire qu'ils ne bougent pas), et puis un vecteur qui est perpendiculaire au plan [11], et dire qu'il change de signe.

Voyons ce qu'il en est pour le plan $x = -z$. Il faut trouver deux vecteurs linéairement indépendants dans ce plan. Prenons par exemple

$$f_1 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad f_2 = \begin{pmatrix}1\\0\\-1\end{pmatrix}. \tag{9.148}$$

Nous avons

$$\begin{aligned} Af_1 &= f_1\\ Af_2 &= f_2. \end{aligned} \tag{9.149}$$

11. Pour le trouver, penser au produit vectoriel.

Afin de trouver un vecteur perpendiculaire au plan, calculons le produit vectoriel :

$$f_3 = f_1 \times f_2 = \begin{vmatrix} e_1 & e_2 & e_3 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{vmatrix} = -e_1 - e_3 = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}. \tag{9.150}$$

Nous avons

$$Af_3 = -f_3. \tag{9.151}$$

Afin de trouver la matrice A, il faut trouver Ae_1, Ae_2 et Ae_3. Pour ce faire, il faut d'abord écrire $\{e_1, e_2, e_3\}$ en fonction de $\{f_1, f_2, f_3\}$. La première des équation (9.148) dit que

$$f_1 = e_2. \tag{9.152}$$

Ensuite, nous avons

$$\begin{aligned} f_2 &= e_1 - e_3 \\ f_3 &= -e_1 - e_3. \end{aligned} \tag{9.153}$$

La somme de ces deux équations donne $-2e_3 = f_2 + f_3$, c'est à dire

$$e_3 = -\frac{f_2 + f_3}{2} \tag{9.154}$$

Et enfin, nous avons

$$e_1 = \frac{f_2 - f_3}{2}. \tag{9.155}$$

Maintenant nous pouvons calculer les images de e_1, e_2 et e_3 en faisant

$$\begin{aligned} Ae_1 &= \frac{Af_2 - Af_3}{2} = \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}, \\ Ae_2 &= Af_1 = f_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \\ Ae_3 &= -\frac{f_2 - f_3}{2} = -\frac{1}{2}\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \tag{9.156}$$

La matrice A s'écrit maintenant en mettant les trois images trouvées en colonnes :

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}. \tag{9.157}$$

9.4.2.2 Symétrie par rapport à une droite

Le principe est exactement le même : il faut trouver trois vecteurs f_1, f_2 et f_3 sur lesquels on connaît l'action de la symétrie. Ensuite il faudra exprimer e_1, e_2 et e_3 en termes de f_1, f_2 et f_3.

Le seul problème est de trouver les trois vecteurs f_i. Le premier est tout trouvé : c'est n'importe quel vecteur sur la droite. Pour les deux autres, il faut un peu ruser parce qu'il faut impérativement qu'ils soient perpendiculaire à la droite. Pour trouver f_2, on peut écrire

$$f_2 = \begin{pmatrix} 1 \\ 0 \\ x \end{pmatrix}, \tag{9.158}$$

et puis fixer le x pour que le produit scalaire de f_2 avec f_1 soit nul. S'il n'y a pas moyen (genre si f_1 a sa troisième composante nulle), essayer avec $\begin{pmatrix} x \\ 1 \\ 0 \end{pmatrix}$. Une fois que f_2 est trouvé (il y a des milliards de choix possibles), trouver f_3 est super facile : prendre le produit vectoriel entre f_1 et f_2.

9.4.2.3 En résumé

La marche à suivre est

(1) Trouver trois vecteurs f_1, f_2 et f_3 sur lesquels on connaît l'action de la symétrie. Typiquement : des vecteurs qui sont sur l'axe ou le plan de symétrie, et puis des perpendiculaires. Pour la perpendiculaire, penser au produit scalaire et au produit vectoriel.

(2) Exprimer la base canonique e_1, e_2 et e_3 en termes de f_1, f_2, f_3.

(3) Trouver Ae_1, Ae_2 et Ae_3 en utilisant leur expression en termes des f_i, et le fait que l'on connaisse l'action de A sur les f_i.

(4) La matrice s'obtient en mettant les images des e_i en colonnes.

9.5 Produit vectoriel

Définition 9.52.
*Soient u et v, deux vecteurs de $\mathbb{R}^3$. Le **produit vectoriel** de u et v est le vecteur $u \times v$ défini par*

$$\begin{aligned} u \times v &= \begin{vmatrix} e_1 & e_2 & e_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \\ &= (u_2v_3 - u_3v_2)e_1 + (u_3v_1 - u_1v_3)e_2 + (u_1v_2 - u_2v_1)e_3 \end{aligned} \tag{9.159}$$

où les vecteurs e_1, e_2 et e_3 sont les vecteurs de la base canonique de $\mathbb{R}^3$.

La notion de produit vectoriel est propre à $\mathbb{R}^3$; il n'y a pas de généralisation simple aux espaces $\mathbb{R}^m$.

Nous n'allons pas nous attarder sur les nombreuses propriétés du produit vectoriel. Les principales sont résumées dans la proposition suivante.

Proposition 9.53.
Si u et v sont des vecteurs de $\mathbb{R}^3$, alors le vecteur $u \times v$ est l'unique vecteur qui est perpendiculaire à u et v en même temps, de norme égal à la surface du parallélogramme construit sur u et v et tel que les vecteurs u, v, $u \times v$ forment une base dextrogyre.

La chose importante à retenir est que le produit vectoriel permet de construire un vecteur simultanément perpendiculaire à deux vecteurs donnés. Le vecteur $u \times v$ est donc linéairement indépendant de u et v. En pratique, si u et v sont déjà linéairement indépendants, alors le produit vectoriel permet de compléter une base de $\mathbb{R}^3$.

À l'aide du produit vectoriel et du produit scalaire, nous construisons le **produit mixte** de trois vecteurs de $\mathbb{R}^3$ par la formule

$$(u \times v) \cdot w = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}. \tag{9.160}$$

Pourquoi nous ne considérons pas la combinaison $(u \cdot v) \times w$?

Proposition 9.54.
Les applications produit scalaire, vectoriel et mixte sont multilinéaires. Spécifiquement, nous avons les propriétés suivantes.

(1) Les applications produit scalaire et vectoriel sont bilinéaires. Le produit mixte est trilinéaire.

(2) Le produit vectoriel est antisymétrique, c'est à dire $u \times v = -v \times u$.

(3) Nous avons $u \times v = 0$ si et seulement si u et v sont colinéaires, c'est à dire si et seulement si l'équation $\alpha u + \beta v = 0$ a une solution différente de la solution triviale $(\alpha, \beta) = (0, 0)$.

(4) Pour tout u et v dans $\mathbb{R}^3$, nous avons

$$\langle u, v\rangle^2 + \|u \times v\|^2 = \|u\|^2\|v\|^2 \tag{9.161}$$

(5) Par rapport à la dérivation, le produit scalaire et vectoriel vérifient une règle de Leibnitz. Soit I un intervalle de $\mathbb{R}$, et si u et u sont dans $C^1(I, \mathbb{R}^3)$, alors

$$\begin{aligned}\frac{d}{dt}\big(u(t) \cdot v(t)\big) &= \big(u'(t) \cdot v(t)\big) + \big(u(t) \cdot v'(t)\big)\\ \frac{d}{dt}\big(u(t) \times v(t)\big) &= \big(u'(t) \times v(t)\big) + \big(u(t) \times v'(t)\big).\end{aligned} \tag{9.162}$$

Les deux formules suivantes, qui mêlent le produit scalaire et le produit vectoriel, sont souvent utiles en analyse vectorielle :

$$\begin{aligned}(u \times v) \cdot w &= u \cdot (v \times w)\\ (u \times v) \times w &= -(v \cdot w)u + (u \cdot w)v\end{aligned} \tag{9.163}$$

pour tout vecteurs u, v et w dans $\mathbb{R}^3$. Nous les admettons sans démonstration. La seconde formule est parfois appelée **formule d'expulsion**.

9.5.1 Produit vectoriel

Une application importante du déterminant 3×3 est qu'il détermine le **produit vectoriel** entre deux vecteurs. Pour cela nous introduisons les vecteurs de base

$$e_x = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad e_y = \begin{pmatrix}0\\1\\0\end{pmatrix}, e_z = \begin{pmatrix}0\\0\\1\end{pmatrix}. \tag{9.164}$$

Ensuite, si v et w sont des vecteurs dans $\mathbb{R}^3$, nous définissons

$$\begin{aligned}\begin{pmatrix}v_x\\v_y\\v_z\end{pmatrix} \times \begin{pmatrix}w_x\\w_y\\w_z\end{pmatrix} = \begin{vmatrix}e_x & e_y & e_z\\ v_x & v_y & v_z\\ w_x & w_y & w_z\end{vmatrix} &= (v_y w_z - w_y vz)e_x\\ &\quad - (v_x w_z - w_x vz)e_y\\ &\quad + (v_x w_y - w_x vy)e_z \in \mathbb{R}^3\end{aligned} \tag{9.165}$$

Ce produit vectoriel peut aussi être écrit sous la forme

$$v \times w = \sum_{i,j,k} \epsilon_{ijk} v_i w_j 1_k. \tag{9.166}$$

où ϵ_{ijk} est défini par $\epsilon_{xyz} = 1$ et ensuite ϵ_{ijk} est 1 ou -1 suivant que la permutation des x, y et z est paire ou impaire. C'est à dire que ϵ_{ijk} est la signature de la permutation qui amène $(1, 2, 3)$ sur (i, j, k).

Proposition 9.55.
Le produit vectoriel est une opération antisymétrique, c'est à dire

$$v \times w = -w \times v. \tag{9.167}$$

En particulier $v \times v = 0$ pour tout vecteur $v \in \mathbb{R}^3$.

Proposition 9.56.
Le produit vectoriel est linéaire. Pour tout vecteurs a, b, c et pour tout nombre α et β nous avons

$$\begin{aligned}a \times (\alpha b + \beta c) &= \alpha(a \times b) + \beta(a \times c)\\ (\alpha a + \beta b) \times c &= \alpha(a \times c) + \beta(b \times c).\end{aligned} \tag{9.168}$$

Les trois vecteurs de base e_x, e_y et e_y ont des produits vectoriels faciles à retenir :

$$\begin{aligned} e_x \times e_y &= e_z \\ e_y \times e_z &= e_x \\ e_z \times e_x &= e_y \end{aligned} \tag{9.169}$$

Exemple 9.57
Calculons le produit vectoriel $v \times w$ avec

$$v = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \quad w = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}. \tag{9.170}$$

Les vecteurs s'écrivent sous la forme $v = 3e_x - e_y + e_z$ et $w = e_x + 2e_y - e_z$. Le produit vectoriel s'écrit

$$\begin{aligned} (3e_x - e_y + e_z) \times (e_x + 2e_y - e_z) &= 6e_x \times e_y - 3e_x \times e_z \\ &\quad - e_y \times e_x + e_y \times e_z \\ &\quad + e_z \times e_x + 2e_z \times e_y \\ &= 6e_z + 3e_y + e_z + e_x + e_y - 2e_x \\ &= -e_x + 4e_y + 7e_z. \end{aligned} \tag{9.171}$$

△

9.5.2 Produit mixte

Si a, b et c sont trois vecteurs, leur **produit mixte** est le nombre $a \cdot (b \times c)$. En écrivant le produit vectoriel sous forme de somme de trois déterminants 2×2, nous avons

$$\begin{aligned} & a \cdot (b \times c) \\ &= (a_1 e_x + a_2 e_y + a_3 e_z) \cdot \left(\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} e_x - \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} e_y + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \right) \\ &= a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \\ &= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \end{aligned} \tag{9.172}$$

Le produit mixte s'écrit donc sous forme d'un déterminant. Nous retenons cette formule :

$$a \cdot (b \times c) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}. \tag{9.173}$$

Un grand intérêt du produit vectoriel est qu'il fournit un vecteur qui est simultanément perpendiculaire aux deux vecteurs donnés.

Proposition 9.58.
Le produit vectoriel $a \times b$ est un vecteur orthogonal à a et b.

Démonstration. Vérifions que $a \perp (a \times b)$. Pour cela, nous calculons $a \cdot (a \times b)$, c'est à dire le produit mixte

$$a \cdot (a \times b) = \begin{vmatrix} a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = 0. \tag{9.174}$$

L'annulation de ce déterminant est due au fait que deux de ses lignes sont égales. □

Ces résultats admettent une intéressante généralisation.

Lemme 9.59.
Soit $X \in \mathbb{R}^n$ ainsi que $v_1, \dots, v_{n-1} \in \mathbb{R}^n$. Alors

(1) Nous avons

$$\det(X, v_1, \dots, v_{n-1}) = X \cdot \det \begin{pmatrix} e_1 & \dots & e_n \\ & v_1 & \\ & \vdots & \\ & v_{n-1} & \end{pmatrix} \tag{9.175}$$

(2) Le vecteur

$$\det \begin{pmatrix} e_1 & \dots & e_n \\ & v_1 & \\ & \vdots & \\ & v_{n-1} & \end{pmatrix} \tag{9.176}$$

est orthogonal à tous les v_i.

Démonstration. Vu que les deux côtés de (9.175) vus comme fonctions de X, sont des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}$, il suffit de vérifier l'égalité sur une base.

Nous posons $\tau_i \colon \mathbb{R}^n \to \mathbb{R}^{n-1}$,

$$\tau_i(v)_k = \begin{cases} v_k & \text{si } k < i \\ v_{k+1} & \text{si } k \geqslant i. \end{cases} \tag{9.177}$$

et nous avons d'une part

$$e_k \cdot \det \begin{pmatrix} e_1 & \dots & e_n \\ & v_1 & \\ & \vdots & \\ & v_{n-1} & \end{pmatrix} = \det \begin{pmatrix} \tau_k v_1 \\ \vdots \\ \tau_k v_{n-1} \end{pmatrix} \tag{9.178}$$

et d'autre part,

$$\det(e_k, v_1, \dots, v_{n-1}) = \det \begin{pmatrix} 0 & & & \\ \vdots & & & \\ 1 & v_1 & \cdots & v_{n-1} \\ \vdots & & & \\ 0 & & & \end{pmatrix} = \det(\tau_k v_1, \dots, \tau_k v_{n-1}). \tag{9.179}$$

La première assertion est démontrée.

En ce qui concerne la seconde, il suffit d'appliquer la première et se souvenir qu'un déterminant est nul lorsque deux lignes sont égales. En effet :

$$v_k \cdot \det \begin{pmatrix} e_1 & \dots & e_n \\ & v_1 & \\ & \vdots & \\ & v_{n-1} & \end{pmatrix} = \det(v_k, v_1, \dots, v_n) = 0. \tag{9.180}$$

□

9.6 Topologie

9.6.1 Boules et sphères

Définition 9.60.
Soit $(V, \|.\|)$, un espace vectoriel normé, $a \in V$ et $r > 0$. Nous allons abondamment nous servir des ensembles suivants :

(1) la ***boule ouverte*** $B(a,r) = \{x \in V \text{ tel que } \|x-a\| < r\}$;

(2) la ***boule fermée*** $\bar{B}(a,r) = \{x \in V \text{ tel que } \|x-a\| \leqslant r\}$;

(3) la ***sphère*** $S(a,r) = \{x \in V \text{ tel que } \|x-a\| = r\}$.

Les différences entre ces trois ensembles sont très importantes. D'abord, les *boules* sont pleines tandis que la *sphère* est creuse. En comparant à une pomme, la boule ouverte serait la pomme «sans la peau», la boule fermée serait «avec la peau» tandis que la sphère serait seulement la peau. Nous avons

$$\bar{B}(a,r) = B(a,r) \cup S(a,r). \tag{9.181}$$

Définition 9.61.
Une partie A de V est dite ***bornée*** *s'il existe un réel R tel que $A \subset B(0_V, R)$.*

Une partie est donc bornée si elle est contenue dans une boule de rayon fini.

Exemple 9.62
Dans $\mathbb{R}$, les boules sont les intervalles ouverts et fermés tandis que la sphère est donnée par les points extrêmes des intervalles :

$$\begin{aligned} B(a,r) &=]a-r, a+r[, \\ \bar{B}(a,r) &= [a-r, a+b], \\ S(a,r) &= \{a-r, a+r\}. \end{aligned} \tag{9.182}$$

△

Exemple 9.63
Si nous considérons $\mathbb{R}^2$, la situation est plus riche parce que nous avons plus de normes. Essayons de voir les sphères de centre $(0,0) \in \mathbb{R}^2$ et de rayon r pour les normes $\|.\|_1$, $\|.\|_2$ et $\|.\|_\infty$.

Pour la norme $\|.\|_1$, la sphère de rayon r est donnée par l'équation

$$|x| + |y| = r. \tag{9.183}$$

Pour la norme $\|.\|_2$, l'équation de la sphère de rayon r est

$$\sqrt{x^2 + y^2} = r, \tag{9.184}$$

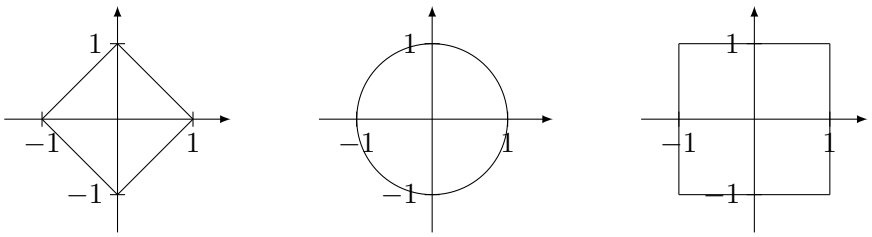
et pour la norme supremum, la sphère de rayon r a pour équation

$$\max\{|x|, |y|\} = r. \tag{9.185}$$

Elles sont dessinées sur la figure 9.1

△

Proposition 9.64.
Soient V un espace vectoriel normé, a dans V et x tel que $d(a,x) = r$, c'est à dire $x \in S(a,r)$. Dans ce cas, toute boule centrée en x contient un point P tel que $d(P,a) > r$ et un point Q tel que $d(Q,a) < r$.



(a) La sphère unité pour la norme $\|.\|_1$ (b) La sphère unité pour la norme $\|.\|_2$ (c) La sphère unité pour la norme $\|.\|_\infty$

FIGURE 9.1 – Les sphères de rayon 1 pour les trois normes classiques.

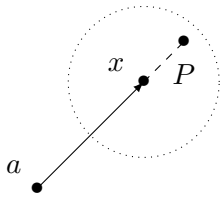


FIGURE 9.2 – Le point P est un peu plus loin que x, en suivant la même droite.

Démonstration. Soit une boule de rayon δ autour de x. Le but est de trouver un point P tel que $d(P,a) > r$ et $d(P,x) < \delta$. Pour cela, nous prenons P sur la même droite que x (en partant de a), mais juste «un peu plus loin» (voir figure 9.2). Plus précisément, nous considérons le point

$$P = x + \frac{v}{N} \tag{9.186}$$

où $v = x - a$ et N est suffisamment grand pour que $d(x,P)$ soit plus petit que δ. Cela est toujours possible parce que

$$d(P,x) = \|P - x\| = \frac{\|v\|}{N} \tag{9.187}$$

peut être rendu aussi petit que l'on veut par un choix approprié de N. Montrons maintenant que $d(a,P) > d(a,x)$:

$$\begin{aligned} d(a,P) &= \|a - x - \frac{v}{N}\| \\ &= \|a - x + \frac{a}{N} - \frac{x}{N}\| \\ &= \|\big(1 + \frac{1}{N}(a-x)\big)\| \\ &> \|a - x\| = d(a,x). \end{aligned} \tag{9.188}$$

Nous laissons en exercice le soin de trouver un point Q tel que $d(Q,a) < r$ et $d(Q,x) < \delta$. □

9.6.2 Ouverts, fermés, intérieur et adhérence

Définition 9.65.
Soit $(V, \|.\|)$ un espace vectoriel normé et A, une partie de V. Un point a est dit ***intérieur*** *à A s'il existe une boule ouverte centrée en a et contenue dans A.*

On appelle ***l'intérieur*** *de A l'ensemble des points qui sont intérieurs à A. Nous notons* $\mathrm{Int}(A)$ *l'intérieur de A.*

Notons que $\mathrm{Int}(A) \subset A$ parce que si $a \in \mathrm{Int}(A)$, nous avons $B(a,r) \subset A$ pour un certain r et en particulier $a \in A$.

Exemple 9.66
Trouver l'intérieur d'un intervalle dans $\mathbb{R}$ consiste à «ouvrir là où c'est fermé».

(1) $\text{Int}\big([0,1[\big) =]0,1[$.
Prouvons d'abord que $]0,1[\subset \text{Int}([0,1[)$. Si $a \in]0,1[$, alors a est strictement supérieur à 0 et strictement inférieur à 1. Dans ce cas, la boule de centre a et de rayon $\frac{\min\{a,1-a\}}{2}$ est contenue dans $]0,1[$ (voir figure 9.3). Cela prouve que a est dans l'intérieur de $[0,1[$.

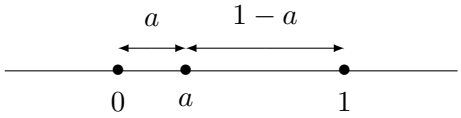


FIGURE 9.3 – Trouver le rayon d'une boule autour de a. Une boule qui serait centrée en a avec un rayon strictement plus petit à la fois de a et de $1-a$ est entièrement contenue dans le segment $]0,1[$.

Prouvons maintenant que $\text{Int}\big([0,1[\big) \subset]0,1[$. Vu que l'intérieur d'un ensemble est inclus à l'ensemble, nous savons déjà que $\text{Int}\big([0,1[\big) \subset [0,1[$. Nous devons donc seulement montrer que 0 n'est pas dans l'intérieur de $[0,1[$. C'est le cas parce que toute boule du type $B(0,r)$ contient le point $-r/2$ qui n'est pas dans $[0,1[$.

(2) $\text{Int}\Big([0,\infty[\Big) =]0,\infty[$.

(3) $\text{Int}\big(]2,3[\big) =]2,3[$.

△

Exemple 9.67
Les intérieurs des boules et sphères sont importantes à savoir.

(1) $\text{Int}\big(B(a,r)\big) = B(a,r)$. Si $x \in B(a,r)$, nous avons $d(a,x) < r$. Alors la boule $B\big(x, r-d(x,a)\big)$ est incluse à $B(a,r)$, et donc x est dans l'intérieur de $B(a,r)$. Conseil : faire un dessin.

(2) $\text{Int}\big(\bar{B}(a,r)\big) = B(a,r)$. Par le point précédent, la boule $B(a,r)$ est certainement dans l'intérieur de la boule fermée. Il reste à montrer que les points de $\bar{B}(a,r)$ qui ne sont pas dans $B(a,r)$ ne sont pas dans l'intérieur. Ces points sont ceux dont la distance à a est *égale* à r. Le résultat découle alors de la proposition 9.64.

(3) $\text{Int}\big(S(a,r)\big) = \varnothing$. Si $x \in S(a,r)$, toute boule centrée en a contient des points qui ne sont pas à distance r de a.
Notez que la sphère est un exemple d'ensemble non vide mais d'intérieur vide.

△

Définition 9.68.
Une partie A de l'espace vectoriel normé $(V, \|.\|)$ est dite ***ouverte*** *si chacun de ses points est intérieur. La partie A est donc ouverte si $A \subset \text{Int}(A)$. Par convention, nous disons que l'ensemble vide $\varnothing$ est ouvert.*

Une partie est dite ***fermée*** *si son complémentaire est ouvert. La partie A est donc fermée si $V \backslash A$ est ouverte.*

Remarque : un ensemble A est ouvert si et seulement si $\text{Int}(A) = A$.

Définition 9.69.
Une partie A de l'espace vectoriel normé V est dite ***compacte*** *si elle est fermée et bornée.*

Nous verrons tout au long de ce cours que les ensembles compacts, et les fonctions définies sur ces ensembles ont de nombreuses propriétés agraables.

Exemple 9.70
En ce qui concerne les intervalles de $\mathbb{R}$,

— $]1, 2[$ est ouvert ;
— $[3, 4]$ est fermé ;
— $[5, 6[$ n'est ni ouvert ni fermé ;

Les intervalles fermés de $\mathbb{R}$ sont toujours compacts. △

Proposition 9.71.
Soit V un espace vectoriel normé.

(1) L'ensemble V lui-même et le vide sont à la fois fermées et ouvertes.
(2) Toute union d'ouverts est ouverte.
(3) Toute intersection finie *d'ouverts est ouverte.*
(4) Le vide et V sont les seules parties de V à être à la fois fermées et ouvertes.

Démonstration. L'ingrédient principal de cette démonstration est que si a est un point d'un ouvert $\mathcal{O}$, alors il existe une boule autour de a contenue dans $\mathcal{O}$ parce que a doit être dans l'intérieur de $\mathcal{O}$.

(1) Nous avons déjà dit que, par définition, l'ensemble vide est ouvert. Cela implique que V lui-même est fermé (parce que son complémentaire est le vide). De plus, V est ouvert parce que toutes les boules sont inclues à V. Le vide est alors fermé (parce que son complémentaire est V).

(2) Soit une famille $(\mathcal{O}_i)_{i\in I}$ d'ouverts[12], et l'union

$$\mathcal{O} = \bigcup_{i\in I} \mathcal{O}_i. \tag{9.189}$$

Soit maintenant $a \in \mathcal{O}$. Nous devons prouver qu'il existe une boule centrée en a entièrement contenue dans $\mathcal{O}$. Étant donné que $a \in \mathcal{O}$, il existe $i \in I$ tel que $a \in \mathcal{O}_i$ (c'est à dire que a est au moins dans un des $\mathcal{O}_i$). Par hypothèse l'ensemble $\mathcal{O}_i$ est ouvert et donc tous ses points (en particulier a) sont intérieurs ; il existe donc une boule $B(a, r)$ centrée en a telle que $B(a, r) \subset \mathcal{O}_i \subset \mathcal{O}$.

(3) Soit une famille finie d'ouverts $(\mathcal{O}_k)_{k\in\{1,\ldots,n\}}$, et $a \in \mathcal{O}$ où

$$\mathcal{O} = \bigcap_{k=1}^{n} \mathcal{O}_k. \tag{9.190}$$

Vu que a appartient à chaque ouvert $\mathcal{O}_k$, nous pouvons trouver, pour chacun de ces ouverts, une boule $B(a, r_k)$ contenue dans $\mathcal{O}_k$. Chacun des r_k est strictement positif, et nous n'en avons qu'un nombre fini, donc le nombre $r = \min\{r_1, \ldots, r_n\}$ est strictement positif. La boule $B(a, r)$ est inclue dans toutes les autres (parce que $B(a, r) \subset B(a, r')$ lorsque $r \leqslant r'$), par conséquent

$$B(a, r) \subset \bigcap_{k=1}^{n} B(a, r_k) \subset \bigcap_{k=1}^{n} \mathcal{O}_k = \mathcal{O}, \tag{9.191}$$

c'est à dire que la boule de rayon r est une boule centrée en a contenue dans $\mathcal{O}$, ce qui fait que a est intérieur à $\mathcal{O}$.

(4) Nous acceptons ce point sans démonstration.

□

La proposition dit que toute intersection *finie* d'ouvert est ouverte. Il est faux de croire que cela se généralise aux intersections infinies, comme le montre l'exemple suivant :

$$\bigcap_{i=1}^{\infty}]-\frac{1}{n}, \frac{1}{n}[= \{0\}. \tag{9.192}$$

12. L'ensemble I avec lequel nous «numérotons» les ouverts $\mathcal{O}_i$ est *quelconque*, c'est à dire qu'il peut être $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}^n$ ou n'importe quel autre ensemble, fini ou infini.

Chacun des ensembles $]-\frac{1}{n},\frac{1}{n}[$ est ouvert, mais le singleton $\{0\}$ est fermé (pourquoi ?).

Nous reportons à la proposition 6.109 la preuve du fait que tout ensemble borné de $\mathbb{R}$ possède un infimum et un supremum.

Définition 9.72.
*L'ensemble des ouverts de V est la **topologie** de V. La topologie dont nous parlons ici est dite **induite** par la norme $\|.\|$ de V (parce que cette norme définit la notion de boule et qu'à son tour la notion de boule définit la notion d'ouverts). Un **voisinage** de a dans V est un ensemble contenant un ouvert contenant a.*

Il existe de nombreuses topologies sur un espace vectoriel donné, mais certaines sont plus fameuses que d'autres. Dans le cas de $V = \mathbb{R}^n$, la topologie **usuelle** est celle induite par la norme euclidienne. Lorsque nous parlons de boules, de fermés, de voisinages ou d'autres notions topologiques (y compris de convergence, voir plus bas) dans $\mathbb{R}^n$, nous sous-entendons toujours la topologie de la norme euclidienne.

Exemple 9.73
Les ensemble suivants sont des voisinages de 3 dans $\mathbb{R}$:
- $]1,5[$;
- $[0,10]$;
- $\mathbb{R}$.

Les ensembles suivants ne sont pas des voisinages de 3 dans $\mathbb{R}$:
- $]1,3[$;
- $]1,3]$;
- $[0,5[\backslash\{3\}$.

△

Proposition 9.74.
Dans un espace vectoriel normé,

(1) toute intersection de fermés est fermée ;

(2) toute union finie *de fermés est fermée.*

Encore une fois, l'hypothèse de finitude de l'intersection est indispensable comme le montre l'exemple suivant :

$$\bigcup_{n=1}^{\infty}[-1+\frac{1}{n},1-\frac{1}{n}] =]-1,1[. \tag{9.193}$$

Chacun des intervalles dont on prend l'union est fermé tandis que l'union est ouverte.

Définition 9.75.
*Soit A, une partie de l'espace vectoriel normé V. Un point $a \in V$ est dit **adhérent** à A dans V si pour tout $\varepsilon > 0$,*

$$B(a,\varepsilon) \cap A \neq \varnothing. \tag{9.194}$$

*Nous notons $\bar{A}$ l'ensemble des points adhérents à a et nous disons que $\bar{A}$ est l'adhérence de A. L'ensemble $\bar{A}$ sera aussi souvent nommé **fermeture** de l'ensemble A.*

Un point peut être adhérent à A sans faire partie de A, et nous avons toujours $A \subset \bar{A}$.

Exemple 9.76
La terminologie «fermeture» de A pour désigner $\bar{A}$ provient de deux origines.

(1) L'ensemble $\bar{A}$ est le plus petit fermé contenant A. Cela signifie que si B est un fermé qui contient A, alors $\bar{A} \subset A$. Nous acceptons cela sans preuve.

(2) Pour les intervalles dans $\mathbb{R}$, trouver $\bar{A}$ revient à fermer les extrémités qui sont ouvertes, comme on en a parlé dans l'exemple 9.70.

△

Exemple 9.77
Dans $\mathbb{R}$, l'infimum et le supremum d'un ensemble sont des points adhérents. En effet si M est le supremum de $A \subset \mathbb{R}$, pour tout ε, il existe un $a \in A$ tel que $a > M - \varepsilon$, tandis que $M > a$. Cela fait que $a \in B(M, \varepsilon)$, et en particulier que pour tout rayon ε, nous avons $B(M, \varepsilon) \cap A \neq \emptyset$.

Le même raisonnement montre que l'infimum est également dans l'adhérence de A. △

Exemple 9.78
Il ne faut pas conclure de l'exemple précédent qu'un point limite ou adhérent est automatiquement un minimum ou un maximum. En effet, si nous regardons l'ensemble formé par les points de la suite $x_n = (-1)^n/n$, le nombre zéro est un point adhérent et une limite, mais pas un infimum ni un maximum. △

Lemme 9.79.
Si B est une partie fermée de V, alors $B = \bar{B}$.

Démonstration. Supposons qu'il existe $a \in \bar{B}$ tel que $a \notin B$. Alors il n'y a pas d'ouverts autour de a qui soit contenu dans $\complement B$. Cela prouve que $\complement B$ n'est pas ouvert, et par conséquent que B n'est pas fermé. Cela est une contradiction qui montre que tout point de $\bar{B}$ doit appartenir à B lorsque B est fermé. □

Exemple 9.80
Au niveau des intervalles dans $\mathbb{R}$, prendre l'adhérence consiste à «fermer là où c'est ouvert». Attention cependant à ne pas fermer l'intervalle en l'infini.

(1) $\overline{[0,2[} = [0,2]$.
(2) $\overline{]3,\infty]} = [3,\infty[$.

△

Proposition 9.81.
Soit V un espace vectoriel normé et $a \in V$. Les trois conditions suivantes sont équivalentes :

(1) $a \in \bar{A}$;
(2) il existe une suite d'éléments x_n dans A qui converge vers a ;
(3) $d(a, A) = 0$.

Notez que dans cette proposition, nous ne supposons pas que a soit dans A.

Proposition 9.82.
Pour toute partie A d'un espace vectoriel normé nous avons

(1) $V \backslash \bar{A} = \operatorname{Int}(V \backslash A)$,
(2) $V \backslash \operatorname{Int}(A) = \overline{V \backslash A}$.

En utilisant les notations du complémentaire (2.1.3), les deux points de la proposition se récrivent

(1) $\complement \bar{A} = \operatorname{Int}(\complement A)$,
(2) $\complement \operatorname{Int}(A) = \overline{\complement A}$.

Démonstration. Nous avons $a \in V \backslash \bar{A}$ si et seulement si $a \notin \bar{A}$. Or ne pas être dans $\bar{A}$ signifie qu'il existe un rayon ε tel que la boule $B(a, \varepsilon)$ n'intersecte pas A. Le fait que la boule $B(a, \varepsilon)$ n'intersecte

pas A est équivalent à dire que $B(a,\varepsilon) \subset V\backslash A$. Or cela est exactement la définition du fait que a est à l'intérieur de $V\backslash A$. Nous avons donc montré que $a \in V\backslash\bar{A}$ si et seulement si $a \in \text{Int}(V\backslash A)$. Cela prouve la première affirmation.

Pour prouver la seconde affirmation, nous appliquons la première au complémentaire de A : $\complement(\overline{\complement A}) = \text{Int}(\complement\complement A)$. En prenant le complémentaire des deux membres nous trouvons successivement

$$\begin{aligned} \complement\complement(\overline{\complement A}) &= \complement\,\text{Int}(\complement\complement A), \\ \overline{\complement A} &= \complement\,\text{Int}(A), \end{aligned} \tag{9.195}$$

ce qu'il fallait démontrer. □

Attention à ne pas confondre $\complement\bar{A}$ et $\overline{\complement A}$. Ces deux ensembles ne sont pas égaux. En effet, en tant que complément d'un fermé, l'ensemble $\complement\bar{A}$ est certainement ouvert, tandis que, en tant que fermeture, l'ensemble $\overline{\complement A}$ est fermé. Pouvez-vous trouver des exemples d'ensembles A tels que $\complement\bar{A} = \overline{\complement A}$?

Proposition 9.83.
Soient A et B deux parties de l'espace vectoriel normé V.

(1) Pour les inclusions, si $A \subset B$, alors $\text{Int}(A) \subset \text{Int}(B)$ et $\bar{A} \subset \bar{B}$.

(2) Pour les unions, $\overline{A \cup B} = \overline{A} \cup \overline{B}$ et $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$.

(3) Pour les intersections, $\text{Int}(A) \cap \text{Int}(B) = \text{Int}(A \cap B)$ et $\text{Int}(A) \cup \text{Int}(B) \subset \text{Int}(A \cup B)$.

Démonstration. (1) Si a est dans l'intérieur de A, il existe une boule autour de a contenue dans A. Cette boule est alors contenue dans B et donc est une boule autour de a contenue dans B, ce qui fait que a est dans l'intérieur de B. Si maintenant a est dans l'adhérence de A, toute boule centrée en a contient un élément de A et donc un élément de B, ce qui prouve que a est dans l'adhérence de B.

(2) Nous avons $A \subset A \cup B$ et donc, en utilisant le premier point, $\bar{A} \subset \overline{A \cup B}$. De la même manière, $\bar{B} \subset \overline{A \cup B}$. En prenant l'union, $\bar{A} \cup \bar{B} \subset \overline{A \cup B}$.
Réciproquement, soit $a \in \overline{A \cup B}$ et montrons que $a \in \bar{A} \cup \bar{B}$. Supposons par l'absurde que a ne soit ni dans $\bar{A}$ ni dans $\bar{B}$. Il existe donc des rayon ε_1 et ε_2 tels que

$$\begin{aligned} B(a,\varepsilon_1) \cap A &= \varnothing, \\ B(a,\varepsilon_2) \cap B &= \varnothing. \end{aligned} \tag{9.196}$$

En prenant $r = \min\{\varepsilon_1, \varepsilon_2\}$, la boule $B(a,r)$ est inclue aux deux boules citées et donc n'intersecte ni A ni B. Donc $a \notin \overline{A \cup B}$, d'où la contradiction.

(3) Si nous appliquons le second point à $\complement A$ et $\complement B$, nous trouvons

$$\overline{\complement A \cup \complement B} = \overline{\complement A} \cup \overline{\complement B}. \tag{9.197}$$

En utilisant les propriétés du lemme 2.12, le membre de gauche devient

$$\overline{\complement A \cup \complement B} = \overline{\complement(A \cap B)} = \complement\,\text{Int}(A \cap B), \tag{9.198}$$

tandis que le membre de droite devient

$$\overline{\complement A} \cup \overline{\complement B} = \complement\,\text{Int}(A) \cup \complement\,\text{Int}(A) = \complement\Big(\text{Int}(A) \cap \text{Int}(B)\Big). \tag{9.199}$$

En égalisant le membre de droite de (9.198) avec celui de (9.199) et en passant au complémentaire nous trouvons

$$\text{Int}(A \cap B) = \text{Int}(A) \cap \text{Int}(B), \tag{9.200}$$

comme annoncé.
La dernière affirmation provient du fait que $\text{Int}(A) \subset \text{Int}(A \cup B)$ et de la propriété équivalente pour B.

□

Remarque 9.84.
Nous avons prouvé que $\overline{A \cap B} \subset \bar{A} \cap \bar{B}$. Il arrive que l'inclusion soit stricte, comme dans l'exemple suivant. Si nous prenons $A = [0, 1]$ et $B =]1, 2]$, nous avons $A \cap B = \varnothing$ et donc $\overline{A \cap B} = \varnothing$. Par contre nous avons $\bar{A} \cap \bar{B} = \{1\}$.

Définition 9.85.
*La **frontière** d'un sous-ensemble A de l'espace vectoriel normé V est l'ensemble des points $a \in V$ tels que*

$$\begin{aligned} B(a, r) \cap A \neq \varnothing, \\ B(a, r) \cap \complement A \neq \varnothing, \end{aligned} \tag{9.201}$$

pour tout rayon r. En d'autres termes, toute boule autour de a contient des points de A et des points de $\complement A$. La frontière de A se note ∂A.

Proposition 9.86.
La frontière d'une partie A d'un espace vectoriel normé V s'exprime sous la forme

$$\partial A = \bar{A} \backslash \operatorname{Int}(A). \tag{9.202}$$

Démonstration. Le fait pour un point a de V d'appartenir à $\bar{A}$ signifie que toute boule centrée en a intersecte A. De la même façon, le fait de ne pas appartenir à $\operatorname{Int}(A)$ signifie que toute boule centrée en a intersecte $\complement A$. □

La description de la frontière donnée par la proposition 9.86 est celle qu'en pratique nous utilisons le plus souvent. Dans certains textes, elle est prise comme définition de la frontière.

Lemme 9.87.
La frontière de A peut également s'exprimer des façons suivantes :

$$\partial A = \bar{A} \cap \complement \operatorname{Int}(A) = \bar{A} \cap \overline{\complement A}, \tag{9.203}$$

Démonstration. En partant de $\partial A = \bar{A} \backslash \operatorname{Int}(A)$, la première égalité est une application de la propriété (4) du lemme 2.12. La seconde égalité est alors la proposition 9.82. □

Exemple 9.88
Dans $\mathbb{R}$, la frontière d'un intervalle est la paire constituée des points extrêmes. En effet

$$\partial[a, b[= \overline{[a, b[} \backslash \operatorname{Int}\big([a, b[\big) = [a, b] \backslash]a, b[= \{a, b\}. \tag{9.204}$$

Toujours dans $\mathbb{R}$ nous avons

$$\partial\mathbb{R} = \bar{\mathbb{R}} \backslash \operatorname{Int}(\mathbb{R}) = \mathbb{R} \backslash \mathbb{R} = \varnothing, \tag{9.205}$$

et

$$\partial\mathbb{Q} = \bar{\mathbb{Q}} \backslash \operatorname{Int}(\mathbb{Q}) = \mathbb{R} \backslash \varnothing = \mathbb{R}. \tag{9.206}$$

△

Exemple 9.89
Dans $\mathbb{R}^n$, nous avons

$$\partial B(a, r) = \partial \bar{B}(a, r) = S(a, r). \tag{9.207}$$

Cela est un boulot pour la proposition 9.64. Si $x \in S(a, r)$ alors tout boule autour de x contient des points à distance strictement plus grande et plus petite que $d(a, x)$, c'est à dire des points dans $B(a, r)$ et hors de $B(a, r)$. Cela prouve que les points de $S(a, r)$ font partie de $\partial B(a, r)$, c'est à dire que $S(a, r) \subset \partial B(a, r)$; et idem pour $\bar{B}(a, r)$.

Pour prouver l'inclusion inverse, soit $x \in \partial B(a, r)$. Vu que toute boule autour de x contient des points intérieurs à $B(a, r)$, pour tout $\epsilon > 0$, $d(a, x) - \epsilon < r$, c'est à dire que $d(a, x) \leqslant r$. De la même manière toute boule autour de x contient des points hors de $B(a, r)$ signifie que pour tout ϵ, $d(a, x) + \epsilon > r$ ou encore que $d(a, x) \geqslant r$. Les deux ensemble implique que $d(a, x) = r$. △

Remarque 9.90.
Il serait toutefois faux de croire que $\partial A = \partial \bar{A}$ pour toute partie A de $\mathbb{R}^n$. En effet si $A = \mathbb{R}\backslash\{0\}$ nous avons $\partial A = \{0\}$ et $\bar{A} = \mathbb{R}$, donc $\partial \bar{A} = \varnothing$.

9.6.3 Point isolé, point d'accumulation

Définition 9.91.
Soit D, une partie de V.

*(1) Un point $a \in D$ est dit **isolé** dans D relativement à V s'il existe un $\varepsilon > 0$ tel que*

$$B(a, \varepsilon) \cap D = \{a\}. \tag{9.208}$$

*(2) Un point $a \in V$ est un **point d'accumulation** de D si pour tout $\varepsilon > 0$,*

$$\Big(B(a, \varepsilon)\backslash\{a\}\Big) \cap D \neq \varnothing. \tag{9.209}$$

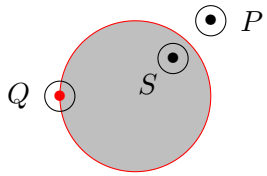


FIGURE 9.4 – L'ensemble décrit par l'équation (9.210). Le point P est un point isolé de D, tandis que les points S et Q sont des points d'accumulation.

Exemple 9.92
Considérons la partie suivante de $\mathbb{R}^2$:

$$D = \{(x, y) \text{ tel que } x^2 + y^2 < 1\} \cup \{(1, 1)\}. \tag{9.210}$$

Comme on peut le voir sur la figure 9.4, le point $P = (1, 1)$ est un point isolé de D parce qu'on peut tracer une boule autour de P sans inclure d'autres points de D que P lui-même. Le point $Q = (-1, 0)$ est un point d'accumulation de D parce que toute boule autour de Q contient des points de D.

Le point S, étant un point intérieur, est un point d'accumulation : toute boule autour de S intersecte D.

Notez cependant que le point Q lui-même n'est pas dans D parce que l'inégalité qui définit D est stricte. △

Remarque 9.93.
À propos de la position des points d'accumulation et des points isolés.

(1) Les points intérieurs sont tous des points d'accumulation.

(2) Les points isolés ne sont jamais intérieurs.

(3) Certains points d'accumulation ne font pas partie de l'ensemble. Par exemple le point 1 est un point d'accumulation de $E =]0, 1[$.

(4) Les points de la frontière sont soit d'accumulation soit isolés.

Exemple 9.94
Tous les points de $\mathbb{R}$ sont des points d'accumulation de $\mathbb{Q}$ parce que dans toute boule autour d'un réel, on peut trouver un nombre rationnel. △

Remarque 9.95.
L'ensemble des points d'accumulation d'un ensemble n'est pas exactement son adhérence. En effet, un point isolé dans A est dans l'adhérence de A, mais n'est pas un point d'accumulation de A.

Lemme 9.96 (Continuité des projections)**.**
Soient deux espaces vectoriels V et W ainsi qu'une fonction continue $f\colon V\times W\to\mathbb{R}$. Alors pour tout $v_0\in V$ la fonction

$$\begin{aligned} f_{v_0}\colon W &\to \mathbb{R} \\ w &\mapsto f(v_0,w) \end{aligned} \tag{9.211}$$

est continue.

9.7 Valeur propre et vecteur propre

Définition 9.97.
*Soit un $\mathbb{K}$-espace vectoriel E et un endomorphisme $A\colon V\to V$. Un **vecteur propre** de A est un vecteur $v\neq 0$ tel que $Av=\lambda v$ pour un certain $\lambda\in\mathbb{K}$. Dans ce cas, λ est la **valeur propre** de v.*

*L'**espace propre** de A pour la valeur λ*[13] *est l'ensemble des vecteurs propres de A pour la valeur propre λ et zéro.*

L'ensemble de valeurs propres de l'endomorphisme u est son **spectre** et est noté $\operatorname{Spec}(u)$.

Remarque 9.98.
Le nombre zéro peut être une valeur propre ; c'est le vecteur zéro qui ne peut pas être vecteur propre. La matrice nulle est une matrice diagonalisable.

Lemme 9.99.
Soit u un endomorphisme et $E_\lambda(u)$ ses espaces propres. La somme des V_λ est directe.

Démonstration. Soit $v_i\in V_{\lambda_i}$ un choix de vecteurs propres de u. Si la somme n'est pas directe, nous pouvons considérer une combinaison linéaire des v_i qui soit nulle :

$$v_1+\cdots+v_p=0. \tag{9.212}$$

Appliquons $(A-\lambda_1\mathbb{1})$ à cette égalité :

$$(\lambda_2-\lambda_1)v_1+\cdots+(\lambda_p-\lambda_1)v_p=0. \tag{9.213}$$

En appliquant encore successivement les opérateurs $(A-\lambda_i\mathbb{1})$ nous réduisons le nombre de termes jusqu'à obtenir $v_p=0$. □

Proposition 9.100 ([97])**.**
Soit E, un espace vectoriel sur un corps infini et $(F_k)_{k=1,\dots,r}$, des sous-espaces vectoriels propres[14] *de E tels que $\bigcup_{i=1}^r F_i=E$. Alors $E=F_k$ pour un certain k.*

Autrement dit, l'union finie de sous-espaces propres ne peut être égal à l'espace complet.

9.8 Polynôme d'endomorphismes

Soit A un anneau commutatif et $\mathbb{K}$, un corps commutatif. L'injection canonique $A\to A[X]$ se prolonge en une injection

$$\mathbb{M}(A)\to\mathbb{M}\big(A[X]\big). \tag{9.214}$$

13. Nous laissons au lecteur le soin de vérifier que c'est bien un sous-espace vectoriel de E.
14. Définition 9.97.

9.8.1 Polynômes d'endomorphismes

Soit $u \in \mathrm{End}(E)$ où E est un $\mathbb{K}$-espace vectoriel. Nous considérons l'application

$$\begin{aligned} \varphi_u \colon \mathbb{K}[X] &\to \mathrm{End}(E) \\ P &\mapsto P(u). \end{aligned} \tag{9.215}$$

L'image de φ_u est un sous-espace vectoriel. En effet si $A = \varphi_u(P)$ et $B = \varphi_u(Q)$, alors $A + B = \varphi_u(P + Q)$ et $\lambda A = (\lambda P)(u)$. En particulier c'est un espace fermé.

Soit u un endomorphisme d'un $\mathbb{K}$-espace vectoriel E et P, un polynôme. Nous disons que P est un polynôme **annulateur** de u si $P(u) = 0$ en tant que endomorphisme de E.

Lemme 9.101.
Si P et Q sont des polynômes dans $\mathbb{K}[X]$ et si u est un endomorphisme d'un $\mathbb{K}$-espace vectoriel E, nous avons

$$(PQ)(u) = P(u) \circ Q(u). \tag{9.216}$$

Démonstration. Si $P = \sum_i a_i X^i$ et $Q = \sum_j b_j X^j$, alors le coefficient de X^k dans PQ est

$$\sum_l a_l b_{k-l}. \tag{9.217}$$

Par conséquent $(PQ)(u)$ contient $\sum_l a_l b_{k-l} u^k$. Par ailleurs $P(u) \circ Q(u)$ est donné par

$$\sum_i a_i u^i \left(\sum_j b_j u^j \right)(x) = \sum_{ij} a_i b_j u^{i+j}(x). \tag{9.218}$$

Le coefficient du terme en u^k est bien le même que celui donné par (9.217). □

Théorème 9.102 (Décomposition des noyaux ou lemme des noyaux)**.**
Soit u un endomorphisme du $\mathbb{K}$-espace vectoriel E. Soit $P \in \mathbb{K}[X]$ un polynôme tel que $P(u) = 0$. Nous supposons que P s'écrive comme le produit $P = P_1 \dots P_n$ de polynômes deux à deux étrangers[15]*. Alors*

$$E = \ker P_1(u) \oplus \dots \oplus \ker P_n(u). \tag{9.219}$$

De plus les projecteurs associés à cette décomposition sont des polynômes en u.

Ce résultat est utilisé pour prouver que toute représentation est décomposable en représentations irréductibles, proposition 17.10 ainsi que pour le théorème 9.143 qui dit que si le polynôme minimal d'un endomorphisme est scindé à racine simple alors il est diagonalisable.

Démonstration. Nous posons

$$Q_i = \prod_{j \neq i} P_i. \tag{9.220}$$

Par le lemme 5.41 ces polynômes sont étrangers entre eux et le théorème de Bézout (théorème 5.40) donne l'existence de polynômes R_i tels que

$$R_1 Q_1 + \dots + R_n Q_n = 1. \tag{9.221}$$

Si nous appliquons cette égalité à u et ensuite à $x \in E$ nous trouvons

$$\sum_{i=1}^{n} (R_i Q_i)(u)(x) = x, \tag{9.222}$$

et en particulier si nous posons $E_i = \mathrm{Image}\big(P_i Q_i(u)\big)$ nous avons

$$E = \sum_{i=1}^{n} E_i. \tag{9.223}$$

15. Définition 4.110.

Cette dernière somme n'est éventuellement pas une somme directe. Si $i \neq j$, alors Q_iQ_j est multiple de P et nous avons, en utilisant le lemme 9.101,

$$(R_iQ_i)(u) \circ (R_jQ_j)(u) = \big(R_iQ_iR_jQ_j\big)(u) = S_{ij}(u) \circ P(u) = 0 \tag{9.224}$$

où S_{ij} est un polynôme.

Nous pouvons voir E comme un $\mathbb{K}$-module et appliquer le théorème 4.31. Les opérateurs $R_iQ_i(u)$ ont l'identité comme somme et sont orthogonaux, et nous avons donc la décomposition en somme directe :

$$E = \bigoplus_{i=1}^{n} R_iQ_i(u)E. \tag{9.225}$$

Afin de terminer la preuve, nous devons montrer que $R_iQ_i(u)E = \ker P_i(u)$. D'abord nous avons

$$P_iR_iQ_i(u) = (R_iP)(u) = R_i(u) \circ P(u) = 0, \tag{9.226}$$

par conséquent $\text{Image}(R_iQ_i(u)) \subset \ker P_i(u)$. Pour obtenir l'inclusion inverse, nous reprenons l'équation (9.222) avec $x \in \ker P_i(u)$. Elle se réduit à

$$(R_iQ_i)(u)x = x. \tag{9.227}$$

Par conséquent $x \in \text{Image}\big(R_iQ_i(u)\big)$. □

Corollaire 9.103.
Soit E, un $\mathbb{K}$-espace vectoriel de dimension finie et f, un endomorphisme semi-simple dont la décomposition du polynôme minimal μ_f en facteurs irréductibles sur $\mathbb{K}[X]$ est $\mu_f = M_1^{\alpha_1} \cdots M_r^{\alpha_r}$. Si F est un sous-espace stable par f, alors

$$F = \bigoplus_{i=1}^{r} \ker M_i^{\alpha_i}(f) \cap F \tag{9.228}$$

Démonstration. Nous posons $E_i = \ker M_i^{\alpha_i}(f)$ et $F_i = E_i \cap F$. Les polynômes $M_i^{\alpha_i}$ sont deux à deux étrangers et $\mu_f(f) = 0$, donc le lemme des noyaux (9.102) s'applique et

$$E = E_1 \oplus \ldots \oplus E_r. \tag{9.229}$$

Nous pouvons décomposer $x \in F$ en termes de cette somme :

$$x = x_1 + \cdots + x_r \tag{9.230}$$

avec $x_i \in E_i$. Toujours selon le lemme des noyaux, les projections sur les espaces E_i sont des polynômes en f. Par conséquent F est stable sous toutes ces projections $\texttt{proj}_i \colon E \to E_i$, et en appliquant $\texttt{proj}_i$ à (9.230), $\texttt{proj}_i(x) = x_i$. Vu que $x \in F$, le membre de gauche est encore dans F et $x_i \in E_i \cap F$. Nous avons donc

$$F \subset \bigoplus_{i=1}^{r} F_i. \tag{9.231}$$

L'inclusion inverse est immédiate parce que $F_i \subset F$ pour chaque i. □

Lemme 9.104.
Si x est un vecteur propre de valeur propre λ pour l'endomorphisme u et si P est un polynôme, alors x est vecteur propre de u pour la valeur propre $P(\lambda)$.

Démonstration. C'est un simple calcul de $P(u)x$ en ayant noté $P(X) = \sum_{k=0}^{n} c_kX^n$:

$$P(u)x = \sum_{k=0}^{n} c_ku^k(x) = \sum_{k=0}^{n} c_k\lambda^k u = P(\lambda)x. \tag{9.232}$$

□

9.8.2 Calcul effectif de l'exponentielle d'une matrice

Nous reprenons l'exemple de [98]. Soit A une matrice dont le polynôme minimum s'écrit

$$P(X) = (X-1)^2(X-2). \tag{9.233}$$

Par le théorème 9.102 de décomposition des noyaux nous avons

$$E = \ker(A-1)^2 \oplus \ker(A-2). \tag{9.234}$$

En suivant les notations de ce théorème nous avons $P_1(X) = (X-1)^2$, $P_2(X) = X-2$ et

$$Q_1(X) = X-2 \tag{9.235a}$$
$$Q_2(X) = (X-1)^2. \tag{9.235b}$$

Les polynômes R_i dont l'existence est assurée par le théorème de Bézout sont

$$\begin{aligned} R_1(X) &= -X \\ R_2(X) &= 1. \end{aligned} \tag{9.236}$$

Nous avons

$$R_1Q_1 + R_2Q_2 = 1. \tag{9.237}$$

Le projecteur p_i sur $\ker P_i$ est R_iQ_i :

$$\begin{aligned} p_1 &= -A(A-2) = \mathtt{proj}_{\ker(u-1)^2} \\ p_2 &= (A-1)^2 = \mathtt{proj}_{\ker(u-2)}\,. \end{aligned} \tag{9.238}$$

Passons maintenant au calcul de l'exponentielle. Nous avons évidemment

$$e^A = e^Ap_1 + e^Ap_2. \tag{9.239}$$

Étant donné que p_1 est le projecteur sur le noyau de $(A-1)^2$, nous avons

$$e^Ap_1 = ee^{A-1}p_1 = ep_1 + e(u-1)1 = ep_1 = -Ae(A-2). \tag{9.240}$$

En effet $e^{A-1}p_1 = \sum_{k=0}^{\infty}(A-1)^k \circ p_1$. De la même façon nous avons

$$e^Ap_2 = e^2e^{A-2}p_2 = e^2p_2 = e^2(A-1)^2. \tag{9.241}$$

Au final,

$$e^A = -Ae(A-2) + e^2(A-1)^2. \tag{9.242}$$

9.8.3 Polynôme minimal et minimal ponctuel

Lemme-définition 9.105.
Soit un endomorphisme $f\colon E \to E$ d'un $\mathbb{K}$-espace vectoriel de dimension finie. Il existe un unique polynôme annulateur normalisé de degré minimum.

*Il est nommé le **polynôme minimal** de f et il est noté μ_f ou simplement μ lorsque la dépendance en f est claire.*

Démonstration. Pour l'unicité, soient P et Q deux polynômes annulateur de f de même degré N et ayant tous deux 1 comme coefficient de x^N. Alors $P-Q$ est de degré $N-1$ tout en étant encore annulateur.

Pour l'existence, les endomorphismes Id, f, f^2, ... ne peuvent pas être tous linéairement indépendants parce que la dimension de $\mathrm{End}(E)$ est finie. Il existe donc un nombre N et des coefficients a_k tels que $\sum_{k=0}^{N} a_kf^k = 0$. Le polynôme $P(X) = \sum_{k=0}^{N} a_kX^k$ est donc annulateur de f.

Une autre façon de le dire est que l'application linéaire $\varphi\colon \mathbb{K}[X] \to \mathrm{End}(E)$ donnée par $\varphi(P) = P(f)$ est un endomorphisme d'un espace vectoriel de dimension infinie vers un espace vectoriel de dimension finie. Il ne peut donc pas être injectif et possède donc un noyau non réduit à zéro. □

Remarque 9.106.
La preuve donnée ci-dessus montre que $\deg(\mu) \leqslant \dim(E)^2$. Comme conséquence du théorème de Caley-Hamilton 9.130 nous verrons qu'en réalité le degré du polynôme minimal est majoré par la dimension de l'espace.

Exemple 9.107(Pas en dimension infinie)
L'endomorphisme de dérivation △

Dans la suite, l'endomorphisme f du $\mathbb{K}$-espace vectoriel E de dimension n est fixé. Pour $x \in E$ nous notons

$$E_x = \{P(f)x \text{ tel que } P \in \mathbb{K}[X]\}. \tag{9.243}$$

Nous considérons le morphisme d'algèbres

$$\begin{aligned} \varphi\colon \mathbb{K}[X] &\to \mathrm{End}(E) \\ P &\mapsto P(f) \end{aligned} \tag{9.244}$$

et si $x \in E$ est donné nous considérons le morphisme de $\mathbb{K}$-espaces vectoriels

$$\begin{aligned} \varphi_x\colon \mathbb{K}[X] &\to E \\ P &\mapsto P(f)x. \end{aligned} \tag{9.245}$$

Les noyaux de ces applications sont des idéaux, entre autres par le lemme 9.101. Ils ont donc un unique générateur unitaire (chacun) par le théorème 5.37. En termes de vocabulaire, l'ensemble

$$\ker(\phi) = \{Q \in \mathbb{K}[X] \text{ tel que } Q(f) = 0\} \tag{9.246}$$

est l'**idéal annulateur** de f et un polynôme Q tel que $Q(f) = 0$ est une polynôme annulateur de f.

Définition 9.108.
Le générateur unitaire de $\ker(\varphi_x)$ *est le* ***polynôme minimal ponctuel*** *de f en x. Il sera noté $\mu_{f,x}$ ou μ_x lorsque la dépendance en f est claire dans le contexte.*

Nous notons μ le générateur unitaire du noyau de φ et μ_x celui de φ_x. Vu que $\mu \in \ker(\varphi_x)$ pour tout x nous avons $\mu_x \mid \mu$ pour tout x.

Exemple 9.109(Pas en dimension infinie)
En dimension infinie, il n'y a pas toujours de polynôme annulateur. Si E est un espace vectoriel de dimension infine ayant une base dénombrable $\{e_i\}_{i\in\mathbb{N}}$ alors l'opérateur donné par $f(e_i) = e_{i+1}$ n'a pas de polynôme annulateur. Même pas ponctuel en quel que point que ce soir.

De même l'opérateur donné par $g(e_1) = 0$ et $g(e_i) = e_{i-1}$ si $i \neq 1$ n'a pas de polynôme annulateur, mais il a un polynôme annulateur ponctuel évident en $x = e_1$. L'exemple 15.56 donnera un habillage à peine subtil à cet exemple. △

Proposition 9.110.
Si P est un polynôme tel que $P(f) = 0$, alors le polynôme minimal μ_f divise P. Autrement dit, le polynôme minimal engendre l'idéal des polynômes annulateurs.

Démonstration. L'ensemble $\ker(\varphi) = \{Q \in \mathbb{K}[X] \text{ tel que } Q(u) = 0\}$ est un idéal par le lemme 9.101. Le polynôme minimal de u est un élément de degré plus bas dans I et par conséquent $I = (\mu_u)$ par le théorème 5.37. Nous concluons que μ_u divise tous les éléments de I. □

La proposition suivante permet de caractériser le polynôme minimal.

Proposition 9.111 ([83])**.**
Soit une application linéaire f sur un $\mathbb{K}$-espace vectoriel. Il existe un unique polynôme unitaire[16] *$P \in \mathbb{K}[X]$ tel que*

16. À mon avis, «unitaire» manque dans [83].

(1) $P(f) = 0$;

(2) l'application

$$\begin{aligned}\varphi\colon \frac{\mathbb{K}[X]}{(P)} &\to \mathrm{End}(E)\\ \bar{Q} &\mapsto Q(f)\end{aligned} \tag{9.247}$$

est injective.

Démonstration. En ce qui concerne l'existence, il existe le polynôme minimal de f qui satisfait les conditions. Pour l'unicité nous travaillons maintenant.

Supposons que l'application (9.247) soit injective. Alors pour tout $Q \in \mathbb{K}[X]$ tel que $Q(f) = 0$ nous avons $\bar{Q} = 0$, c'est à dire $Q = PR$ pour un certain $R \in \mathbb{K}[X]$. Autrement dit : P est un générateur unitaire de l'idéal annulateur de f. Le théorème 5.37(3) nous dit alors que $P = \mu$ parce que μ est également générateur unitaire. □

Lemme 9.112 ([99])**.**
Soit $f\colon E \to E$ un endomorphisme de l'espace vectoriel E. Il existe un élément $x \in E$ tel que $\mu_{f,x} = \mu_f$.

Démonstration. Soit une décomposition en irréductibles du polynôme minimal $\mu = P_1^{\alpha_1} \dots P_r^{\alpha_r}$. Nous notons $E_i = \ker\big(P_i^{\alpha_i}(f)\big)$. Les polynômes P_i sont étrangers deux à deux (un diviseur commun aurait a fortiori été un diviseur et aurait contredit l'irréductibilité). Le lemme des noyaux 9.102 nous donne la somme directe

$$E = \bigoplus_{i=1}^{r} \ker\big(P_i^{\alpha_i}(f)\big). \tag{9.248}$$

Si $x_i \in E_i$ alors μ_{x_i} est une puissance de P_i. En effet $\mu_{x_i} \mid \mu$ et est donc un produit des puissances des P_j. Or si $(QP_j)(f)x_i = 0$ alors $(P_jQ)(f)x_i = 0$, ce qui donne $Q(f)x_i \in E_j \cap E_i = \{0\}$. Donc μ_{x_i} n'est pas de la forme QP_j pour $j \neq i$. Nous en déduisons que μ_{x_i} est une puissance de P_i dès que $x_i \in E_i$. Nous choisissons $x_i \in E_i$ tel que $\mu_{x_i} = P_i^{\alpha_i}$.

Nous posons enfin $a = x_1 + \dots + x_r$; par définition du polynôme annulateur μ_a, nous avons

$$0 = \mu_a(f)a = \mu_a(f)x_1 + \dots + \mu_a(f)x_r. \tag{9.249}$$

Mais $m_a(f)x_j \in E_i$, et la somme des E_j est directe, donc l'annulation de la somme (9.249) implique l'annulation de chacun des termes : $\mu_a(f)x_i = 0$ pour tout i. Cela prouve que $\mu_{x_i} \mid \mu_a$. Mais comme les μ_{x_i} sont premiers deux à deux (parce que ce sont les $P_i^{\alpha_i}$), nous avons que le produit divise encore μ_a :

$$\prod_{i=1}^{r} \mu_{x_i} \mid \mu_a, \tag{9.250}$$

c'est à dire $\mu \mid \mu_a$. Comme nous avons aussi $\mu_a \mid \mu$, nous déduisons $\mu_a = \mu$. □

Définition 9.113 (Matrices, endomorphismes et vecteurs cycliques)**.**
Une matrice est ***cyclique*** *si elle est semblable à une matrice compagnon. Un endomorphisme $f\colon E \to E$ est* ***cyclique*** *s'il existe un vecteur $x \in E$ tel que $\{f^k(x)\}_{k=0,\dots,n-1}$ est une base de E. Un vecteur ayant cette propriété est un* ***vecteur cyclique*** *pour f.*

Lemme 9.114.
Soit E un espace vectoriel de dimension finie et un endomorphisme cyclique[17] *f de E. Soit un vecteur cyclique v de f, alors le polynôme minimal de f est égal au polynôme minimal de f au point v : $\mu_f = \mu_{f,v}$.*

Démonstration. Montrons que $\mu_{f,v}$ est un polynôme annulateur de f, ce qui prouvera que μ_f divise $\mu_{f,v}$ par la proposition 9.110. Étant donné que v est cyclique, tout élément de E s'écrit sous la

17. Voir la définition 9.113.

forme $x = Q(f)v$. Prenons un polynôme P annulateur de f en v : $P(f)v = 0$. Nous montrons que P est alors un polynôme annulateur de f. En effet, nous avons

$$P(f)x = \big(P(f) \circ Q(f)\big)v = \big(Q(f) \circ P(f)\big)v = 0 \tag{9.251}$$

où nous avons utilisé le lemme 9.101. □

Lemme 9.115 ([99])**.**
Soit $a \in E$ tel que $\mu_a = \mu$. Alors E_a est un sous-espace stable pour f pour lequel il existe un supplémentaire stable.

Démonstration. Soit $l = \deg(\mu) = \deg(\mu_a)$. L'espace E_a étant engendré par les $f^k(a)$ nous savons que $e_1 = a$, $e_2 = f(a)$,..., $e_l = k^{l-1}(a)$ forment une base de E_a. Nous pouvons la compléter en une base $\{e_1, \dots, e_n\}$ de E. Et nous posons [18]

$$\begin{aligned} G &= \{x \in E \text{ tel que } e_l^*\big(f^k(x)\big) = 0 \, \forall k \geqslant 0\} && (9.252a)\\ &= \bigcap_{k\geqslant 0} \ker\{e_l^* \circ f^k\} && (9.252b)\\ &= \bigcap_{k=0}^{l-1} \ker(e_l^* \circ f^k). && (9.252c) \end{aligned}$$

La dernière égalité est due au fait que l soit le degré de μ. Du coup f^l est une combinaison linéaire des f^i avec $i \leqslant l-1$.

Nous avons $f(G) \subset G$ et de plus $E_a \cap G = \{0\}$ parce qu'un élément de E_a est une combinaison linéaire d'éléments de la forme $f^j(a)$ $(j \leqslant l)$. Après application de f^{l-j}, ces éléments obtiennent une composante $f^l(a) = e_l$. De plus G est un sous-espace vectoriel du fait que $e_l^* \circ f^i$ est une application linéaire.

Montrons enfin que $\dim(G) = n - l$. Pour cela nous remarquons que G est une intersection d'hyperplans, et nous montrons que les équations définissant ces hyperplans sont linéairement indépendantes. Soit donc

$$\sum_{j=0}^{l-1} \lambda_j\big(e_l^* \circ f^j\big) = 0 \tag{9.253}$$

et montrons que $\lambda_j = 0$ pour tout j est l'unique solution. Soit $x \in E$ et appliquons l'opération (9.253) au vecteur $f^i(x)$; le résultat est zéro :

$$0 = \sum_{j=0}^{l-1} \lambda_j(e_l^* \circ f^i \circ f^j) = (e_l^* \circ f^i)P(u) \tag{9.254}$$

où nous avons posé $P(X) = \sum_{j=0}^{l-1} \lambda_j X^j$. Appliquons cela à a : pour tout i nous avons

$$(e_l^* \circ f^i)\big(P(f)a\big) = 0. \tag{9.255}$$

Mais par définition de E_a, l'élément $P(f)a$ est dans E_a. Nous en déduisons que

$$P(f)a \in G \cap E_a = \{0\}, \tag{9.256}$$

c'est à dire que P est un polynôme annulateur de a. Mais P est de degré $l-1$ alors que le polynôme minimal de a est de degré l. Par conséquent $P = 0$ et $\lambda_j = 0$ pour tout j. □

Définition 9.116.
*Un endomorphisme d'un espace vectoriel est **semi-simple** si tout sous-espace stable par u possède un supplémentaire stable.*

18. ici, comme presque partout, e_l^* est le dual de e_l, c'est à dire l'application linéaire sur E donnée par $e_l^*(e_i) = \delta_{li}$.

Lemme 9.117.
Si le polynôme minimal d'un endomorphisme est irréductible, alors il est semi-simple[19].

Démonstration. Soit f, un endomorphisme dont le polynôme minimal est irréductible et F, un sous-espace stable par f. Nous devons en trouver un supplémentaire stable. Si $F = E$, il n'y a pas de problèmes. Sinon nous considérons $u_1 \in E\backslash F$ et

$$E_{u_1} = \{P(f)u_1 \text{ tel que } P \in \mathbb{K}[X]\}, \tag{9.257}$$

qui est un espace stable par f.

Montrons que $E_{u_1} \cap F = \{0\}$. Pour cela nous regardons l'idéal

$$I_{u_1} = \{P \in \mathbb{K}[X] \text{ tel que } P(f)u_1 = 0\}. \tag{9.258}$$

Cela est un idéal non réduit à $\{0\}$ parce que le polynôme minimal de f par exemple est dans I_{u_1}. Soit P_{u_1} un générateur unitaire de I_{u_1}. Étant donné que $\mu_f \in I_{u_1}$, nous avons que P_{u_1} divise μ_f et donc $P_{u_1} = \mu_f$ parce que μ_f est irréductible par hypothèse.

Soit $y \in E_{u_1} \cap F$. Par définition il existe $P \in \mathbb{K}[X]$ tel que $y = P(f)u_1$ et si $y \neq 0$, ce la signifie que $P \notin I_{u_1}$, c'est à dire que P_{u_1} ne divise pas P. Étant donné que P_{u_1} est irréductible cela implique que P_{u_1} et P sont premiers entre eux (ils n'ont pas d'autre pgcd que 1).

Nous utilisons maintenant Bézout (théorème 5.40) qui nous donne $A, B \in \mathbb{K}[X]$ tels que

$$AP + BP_{u_1} = 1. \tag{9.259}$$

Nous appliquons cette égalité à f et puis à u_1 :

$$u_1 = A(f) \circ \underbrace{P(f)u_1}_{=y} + B(f) \circ \underbrace{P_{u_1}(u_1)}_{=0} = A(f)y. \tag{9.260}$$

Mais $y \in F$, donc $A(f)y \in F$. Nous aurions donc $u_1 \in F$, ce qui est impossible par choix. Nous avons maintenant que l'espace $E_{u_1} \oplus F$ est stable sous f. Si cet espace est E alors nous arrêtons. Sinon nous reprenons le raisonnement avec $E_{u_1} \oplus F$ en guise de F et en prenant $u_2 \in E\backslash(E_{u_1} \oplus F)$. Étant donné que E est de dimension finie, ce procédé s'arrête à un certain moment et nous aurons

$$E = F \oplus E_{u_1} \oplus \ldots \oplus E_{u_k} \tag{9.261}$$

où chacun des E_{u_i} sont stables. □

Théorème 9.118.
Un endomorphisme est semi-simple si et seulement si son polynôme minimal est produit de polynômes irréductibles distincts deux à deux.

Démonstration. Supposons que f soit semi-simple et que son polynôme minimal soit donné par $\mu_f = M_1^{\alpha_1} \ldots M_r^{\alpha_r}$ où les M_i sont des polynômes irréductibles deux à deux distincts. Nous devons montrer que $\alpha_i = 1$ pour tout i. Soit i tel que $\alpha_i \geqslant 1$ et $N \in \mathbb{K}[X]$ tel que $\mu_f = M^2N$ où l'on a noté $M = M_i$. Nous étudions l'espace

$$F = \ker M(f) \tag{9.262}$$

qui est stable par f, et qui possède donc un supplémentaire S également stable par f. Nous allons montrer que MN est un polynôme annulateur de f.

D'abord nous prenons $x \in S$. Étant donné que F est le noyau de $M(f)$,

$$M(f)\big(MN(f)x\big) = \mu_f(f)x = 0, \tag{9.263}$$

19. Définition 9.116.

ce qui signifie que $MN(f)x \in F$. Mais vu que S est stable par f nous avons aussi que $MN(f)x \in S$. Finalement $MN(f)x \in F \cap S = \{0\}$. Autrement dit, $MN(f)$ s'annule sur S.

Prenons maintenant $y \in F$. Nous avons

$$MN(f) = N(f)\big(M(f)y\big) = 0 \tag{9.264}$$

parce que $y \in F = \ker M(f)$.

Nous avons prouvé que $MN(f)$ s'annule partout et donc que $MN(f)$ est un polynôme annulateur de f, ce qui contredit la minimalité de $\mu_f = M^2N$.

Nous passons au sens inverse. Soit $m_f = M_1 \dots M_r$ une décomposition du polynôme minimal de l'endomorphisme f en irréductibles distincts deux à deux. Soit F un sous-espace vectoriel stable par f. Nous notons

$$E_i = \ker(M_i(f)) \tag{9.265}$$

et $f_i = f|_{E_i}$. Par le lemme 9.103 nous avons

$$F = \bigoplus_{i=1}^{r}(F \cap E_i). \tag{9.266}$$

Les espaces E_i sont stables par f et étant donné que M_i est irréductible, il est le polynôme minimal de f_i. En effet, M_i est annulateur de f_i, ce qui montre que le minimal de f_i divise M_i. Mais M_i étant irréductible, M_i est le polynôme minimal. Étant donné que $\mu_{f_i} = M_i$, l'endomorphisme f_i est semi-simple par le lemme 9.117.

L'espace $F \cap E_i$ étant stable par l'endomorphisme semi-simple f_i, il possède un supplémentaire stable que nous notons S_i :

$$E_i = S_i \oplus (F \cap E_i). \tag{9.267}$$

Étant donné que sur chaque S_i nous avons $f|_{S_i} = f_i$, l'espace $S = S_1 \oplus \dots \oplus S_r$ est stable par f. Du coup nous avons

$$\begin{aligned}
E &= E_1 \oplus \dots \oplus E_r && \text{(9.268a)}\\
&= \big(S_1 \oplus (F \cap E_1)\big) \oplus \dots \oplus \big(S_r \oplus (F \cap E_r)\big) && \text{(9.268b)}\\
&= \big(\bigoplus_{i=1}^{r} S_i\big) \oplus \big(\bigoplus_{i=1}^{r} F \cap E_i\big) && \text{(9.268c)}\\
&= S \oplus F, && \text{(9.268d)}
\end{aligned}$$

ce qui montre que F a bien un supplémentaire stable par f et donc que f est semi-simple. □

Exemple 9.119(L'espace engendré par $\mathbb{1}$, A, A^2,...)
Soit A une matrice, et

$$V = \text{Span}\{A^k \text{ tel que } k \in \mathbb{N}\}. \tag{9.269}$$

Nous montrons que $\dim(V)$ est le degré du polynôme minimal de A.

D'abord l'idéal annulateur de A est engendré par le polynôme minimal[20] que nous notons $\mu = \sum_{k=0}^{p} a_k X^k$. La partie $\{\mathbb{1}, \dots, A^{p-1}\}$ est libre parce qu'une combinaison linéaire nulle de cela serait un polynôme annulateur en A de degré plus petit que p. Donc $\dim(V) \geqslant p$.

La partie $\{\mathbb{1}, A, \dots, A^p\}$ est liée à cause du polynôme minimal. Isoler A^p dans $\mu(A) = 0$ donne un polynôme f de degré $p-1$ tel que $A^p = f(A)$.

Nous allons montrer à présent que la famille $\{\mathbb{1}, A, \dots, A^{p-1}\}$ est génératrice (alors $\dim(V) \leqslant p$). Soit un entier $q \geqslant p$et de division euclidienne[21] $np + r = q$ avec $r < p$. Nous avons $A^q = A^{np}A^r$. D'une part

$$A^{np} = (A^p)^n = f(A)^n \tag{9.270}$$

20. Proposition 9.110.
21. Théorème 3.25.

est de degré $n(p-1)$. Par conséquent

$$A^q = f(A)^n A^r \tag{9.271}$$

qui est de degré $n(p-1)+r = q-n$. Autrement dit il existe un polynôme g_1 de degré $q-n$ tel que $A^q = g_1(A)$. Si $q-n > p-1$ alors nous pouvons recommencer et obtenir un polynôme g_2 de degré strictement inférieur à celui de g_1 tel que $A^q = g_2(A)$. Au bout du compte, il existe un polynôme g de degré au maximum $p-1$ tel que $A^q = g(A)$. Cela prouve que la partie $\{\mathbb{1}, A, \dots, A^{p-1}\}$ est génératrice de V.

La dimension de V est donc p, le degré du polynôme minimal. △

Proposition 9.120.
Soit f un endomorphisme d'un espace vectoriel de dimension finie. Nous avons l'isomorphisme d'espace vectoriel

$$\mathbb{K}[f] \simeq \frac{\mathbb{K}[X]}{(\mu_f)} \tag{9.272}$$

La dimension en est $\deg(\mu_f)$.

Démonstration. Notons avant de commencer que (μ) est l'idéal engendré par μ. Les classes dont il est question dans le quotient $\mathbb{K}[X]/(\mu)$ sont

$$\bar{P} = \{P + S\mu\}_{S\in\mathbb{K}[X]}. \tag{9.273}$$

Nous allons montrer que l'application suivante fournit l'isomorphisme :

$$\begin{aligned} \psi\colon \frac{\mathbb{K}[X]}{(\mu)} &\to \mathbb{K}[f] \\ \bar{P} &\mapsto P(f). \end{aligned} \tag{9.274}$$

ψ est bien définie Si $Q \in \bar{P}$ alors $Q = P + S\mu$ pour un certain $S \in \mathbb{K}[X]$. Du coup nous avons

$$\psi(\bar{Q}) = P(f) + (S\mu)(f). \tag{9.275}$$

Mais $\mu(f) = 0$ donc le deuxième terme est nul. Donc $\psi(\bar{P})$ est bien définit.

Injectif Si $\psi(\bar{P}) = 0$ nous avons $P(f) = 0$, ce qui signifie que $P = S\mu$ pour un polynôme S.
Par conséquent $P \in (\mu)$ et donc $\bar{P} = 0$.

Surjectif Soit $P \in \mathbb{K}[X]$. L'élément $P(f)$ de $\mathbb{K}[f]$ est dans l'image de ψ parce que c'est $\psi(\bar{P})$.

En ce qui concerne la dimension, le corollaire 5.38 en parle déjà : une base est donné par les projections de $1, X, \dots, X^{\deg(\mu_a)-1}$. □

9.8.4 Polynôme caractéristique

Définition 9.121.
Soit un anneau commutatif A. Si $u \in \mathbb{M}_n(A)$, nous définissons le ***polynôme caractéristique de*** *u :*

$$\chi_u(X) = \det(u - X\mathbb{1}_n). \tag{9.276}$$

Nous définissons de même le polynôme caractéristique d'un endomorphisme $u\colon E \to E$.

Remarque 9.122.
Quelque remarques à propos du signe[22] .

- Certains auteurs définissent le polynôme caractéristique par $\det(X-u)$ au lieu de $\det(u-X)$.
- Wikipédia francophone prend la définition $\det(X-u)$ (donc inverse de la notre). Allez lire la page de discussion.

22. Attention : je crois qu'il y a des incohérences dans le Frido à propos de ce choix

— Sur les wikipédias d'autre langues, ça varie.
— Un avantage de $\det(u - X)$ est que $\det(u) = \chi_u(0)$.
— Un avantage de $\det(X - u)$ est qu'il est unitaire.

Lemme 9.123.
Le polynôme caractéristique χ_u est unitaire en dimension paire et a pour degré la dimension de l'espace vectoriel E..

Théorème 9.124.
Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n et un endomorphisme $u \in \mathrm{End}(E)$. Alors

(1) Le polynôme caractéristique divise $(\mu_u)^n$ dans $\mathbb{K}[X]$.

(2) Les polynômes caractéristiques et minimaux ont mêmes facteurs irréductibles dans $\mathbb{K}[X]$.

(3) Les polynômes caractéristiques et minimaux ont mêmes racines dans $\mathbb{K}[X]$.

(4) Le polynôme caractéristique est scindé si et seulement si le polynôme minimal est scindé.

Théorème 9.125.
Soit $u \in \mathrm{End}(E)$ et $\lambda \in \mathbb{K}$. Les conditions suivantes sont équivalentes

(1) $\lambda \in \mathrm{Spec}(u)$

(2) $\chi_u(\lambda) = 0$

(3) $\mu_u(\lambda) = 0$.

Démonstration. (1) $\Leftrightarrow$ (2). Dire que λ est dans le spectre de u signifie que l'opérateur $u - \lambda\mathbb{1}$ n'est pas inversible, ce qui est équivalent à dire que $\det(u - \lambda\mathbb{1})$ est nul par la proposition 9.10(1) ou encore que λ est une racine du polynôme caractéristique de u.

(2) $\Leftrightarrow$ (3). Cela est une application directe du théorème 9.124 qui précise que le polynôme caractéristique a les mêmes racines dans $\mathbb{K}$ que le polynôme minimal. □

Définition 9.126.
*Si $\lambda \in \mathbb{K}$ est une racine de χ_u, l'ordre de l'annulation est la **multiplicité algébrique** de la valeur propre λ de u. À ne pas confondre avec la **multiplicité géométrique** qui sera la dimension de l'espace propre.*

Exemple 9.127

Sur $\mathbb{R}^2$, nous considérons la matrice $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ qui a pour polynôme caractéristique[23] le polynôme $\chi_A = (X - 1)^2$. Le nombre $\lambda = 1$ est une racine double de ce polynôme, et pourtant il n'y a qu'une seule dimension d'espace propre :

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{9.277}$$

entraine $x = 0$.

Ici la multiplicité algébrique est différente de la multiplicité géométrique. △

La proposition suivante donne une utilisation amusante de la notion de polynôme caractéristique[24].

Proposition 9.128 ([100])**.**
Soit un espace vectoriel V de dimension finie pour lequel il existe un endomorphisme $f : V \to V$ tel que $(f \circ f)(v) = -v$ pour tout $v \in V$. Alors la dimension de V est paire.

23. Définition 9.121.
24. Définition 9.121.

Démonstration. Cherchons les valeurs propres de f en résolvant l'équation $f(v) = \lambda v$. Nous appliquons f à cette égalité :

$$-v = \lambda f(v) = \lambda^2 v. \tag{9.278}$$

Donc λ ne peut pas être réel. Nous avons montré que f n'a pas de valeurs propres réelles. Or le polynôme caractéristique de f est de degré égal à la dimension. Si la dimension est impaire, le polynôme caractéristique est de degré impair, et possède donc une racine réelle. Autrement dit, l'absence de racines réelles au polynôme caractéristique indique une dimension paire. □

Une autre preuve possible est d'utiliser le déterminant : si la dimension de V est n nous avons :

$$\det(f^2) = \det(-\mathrm{Id}) = (-1)^n. \tag{9.279}$$

Donc $(-1)^n$ est positif, ce qui montre que n est pair.

Proposition 9.129 ([97])**.**
Soit f, un endomorphisme de E et $x \in E$. Alors

(1) L'espace $E_{f,x}$ est stable par f.

(2) L'espace $E_{f,x}$ est de dimension

$$p_{f,x} = \dim E_{f,x} = \deg(\mu_{f,x}) \tag{9.280}$$

où $\mu_{f,x}$ est le générateur unitaire de $I_{f,x}$.

(3) Le polynôme caractéristique de $f|_{E_{f,x}}$ est $\mu_{f,x}$.

(4) Nous avons

$$\chi_{f|_{E_{f,x}}}(f)x = \mu_{f,x}(f)x = 0. \tag{9.281}$$

Démonstration. Le fait que $E_{f,x}$ soit stable par f est classique. Le point (4) est un une application du point (3). Les deux gros morceaux sont donc les points (2) et (3).

Étant donné que $\mu_{f,x}$ est de degré minimal dans $I_{f,x}$, l'ensemble

$$B = \{f^k(x) \text{ tel que } 0 \leqslant k \leqslant p_{f,x} - 1\} \tag{9.282}$$

est libre. En effet une combinaison nulle des vecteurs de B donnerait un polynôme en f de degré inférieur à $p_{f,x}$ annulant x. Nous écrivons

$$\mu_{f,x}(X) = X^{p_{f,x}} - \sum_{i=0}^{p_{f,x}-1} a_i X^k. \tag{9.283}$$

Étant donné que $\mu_{f,x}(f)x = 0$ et que la somme du membre de droite est dans $\mathrm{Span}(B)$, nous avons $f^{p_{f,x}}(x) \in \mathrm{Span}(B)$. Nous prouvons par récurrence que $f^{p_{f,x}+k}(x) \in \mathrm{Span}(B)$. En effet en appliquant f^k à l'égalité

$$0 = f^{p_{f,x}}(x) - \sum_{i=0}^{p_{f,x}-1} a_i f^i(x) \tag{9.284}$$

nous trouvons

$$f^{p_{f,x}+k}(x) = \sum_{i=0}^{p_{f,x}-1} a_i f^{i+k}(x), \tag{9.285}$$

alors que par hypothèse de récurrence le membre de droite est dans $\mathrm{Span}(B)$. L'ensemble B est alors générateur de $E_{f,x}$ et donc une base d'icelui. Nous avons donc bien $\dim(E_{f,x}) = p_{f,x}$.

Nous montrons maintenant que $\mu_{f,x}$ est annulateur de f au point x. Nous savons que

$$\mu_{f,x}(f)x = 0. \tag{9.286}$$

En y appliquant f^k et en profitant de la commutativité des polynômes sur les endomorphismes (proposition 9.101), nous avons

$$0 = f^k\big(\mu_{f,x}(f)x\big) = \mu_{f,x}(f)f^k(x), \tag{9.287}$$

de telle sorte que $\mu_{f,x}(f)$ est nul sur B et donc est nul sur $E_{f,x}$. Autrement dit,

$$\mu_{f,x}\big(f|_{E_{f,x}}\big) = 0. \tag{9.288}$$

Montrons que $\mu_{f,x}$ est même minimal pour $f|_{E_{f,x}}$. Sot Q, un polynôme non nul de degré $p_{f,x} - 1$ annulant $f|_{E_{f,x}}$. En particulier $Q(f)x = 0$, alors qu'une telle relation signifierait que B est un système lié, alors que nous avons montré que c'était un système libre. Nous concluons que $\mu_{f,x}$ est le polynôme minimal de $f|_{E_{f,x}}$. □

Cette histoire de densité permet de donner une démonstration alternative du théorème de Cayley-Hamilton.

Théorème 9.130 (Cayley-Hamlilton)**.**
Le polynôme caractéristique est un polynôme annulateur.

Une démonstration plus simple via la densité des diagonalisables est donnée en théorème 12.339.

Démonstration. Nous devons prouver que $\chi_f(f)x = 0$ pour tout $x \in E$. Pour cela nous nous fixons un $x \in E$, nous considérons l'espace $E_{f,x}$ et $\chi_{f,x}$, le polynôme caractéristique de $f|_{E_{f,x}}$. Étant donné que $E_{f,x}$ est stable par f, le polynôme caractéristique de $f|_{E_{j,x}}$ divise χ_f, c'est à dire qu'il existe un polynôme Q_x tel que

$$\chi_f = Q_x\chi_{f,x}, \tag{9.289}$$

et donc aussi

$$\chi_f(f)x = Q_x(f)\big(\chi_{f,x}(f)x\big) = 0 \tag{9.290}$$

parce que la proposition 9.129 nous indique que $\chi_{f,x}$ est un polynôme annulateur de $f|_{E_{f,x}}$. □

Corollaire 9.131.
Le degré du polynôme minimal est majoré par la dimension de l'espace.

Démonstration. Le polynôme minimal engendre l'idéal des polynôme annulateurs (proposition 9.110), et divise donc le polynôme caractéristique. Or le degré du polynôme caractéristique est la dimension de l'espace par le lemme 9.123. □

Exemple 9.132(Calcul de l'inverse d'un endomorphisme)
Le polynôme de Cayley-Hamilton donne un moyen de calculer l'inverse d'un endomorphisme inversible pourvu que l'on sache son polynôme caractéristique. En effet, supposons que

$$\chi_f(X) = \sum_{k=0}^{n} a_k X^k. \tag{9.291}$$

Nous aurons alors

$$0 = \chi_f(f) = \sum_{k=0}^{n} a_k f^k. \tag{9.292}$$

Nous appliquons f^{-1} à cette dernière égalité en sachant que $f^{-1}(0) = 0$:

$$0 = a_0 f^{-1} + \sum_{k=1}^{n} a_k f^{k-1}, \tag{9.293}$$

et donc

$$u^{-1} = -\frac{1}{\det(f)} \sum_{k=1}^{n} a_k f^{k-1} \tag{9.294}$$

où nous avons utilisé le fait que $a_0 = \chi_f(0) = \det(f)$. △

Proposition 9.133.
Si $(X-z)^l$ ($l \geqslant 1$) est la plus grande puissance de $(X-z)$ dans le polynôme caractéristique d'un endomorphisme u alors

$$1 \leqslant \dim(E_e) \leqslant l. \tag{9.295}$$

C'est à dire que nous avons au moins un vecteur propre pour chaque racine du polynôme caractéristique.

Démonstration. Si $(X-z)$ divise χ_u alors en posant $\chi_u = (X-z)P(X)$ nous avons

$$\det(u - X\mathbb{1}) = (X-z)P(X), \tag{9.296}$$

ce qui, évalué en $X = z$, donne $\det(u - z\mathbb{1}) = 0$. L'annulation du déterminant étant équivalente à l'existence d'un noyau non trivial, nous avons $v \neq 0$ dans E tel que $(u - z\mathbb{1})v = 0$. Cela donne $u(v) = zv$ et donc que v est vecteur propre de u pour la valeur propre z. Donc aussi $\dim(E_z) \geqslant 1$.

Si $\dim(E_z) = k$ alors le théorème de la base incomplète 7.10 nous permet d'écrire une base de E dont les k premiers vecteurs forment une base de E_z. Dans cette base, la matrice de u est de la forme

$$\begin{pmatrix} z & & & * \\ & \ddots & & \vdots \\ & & z & * \\ & & & * \end{pmatrix} \tag{9.297}$$

où les étoiles représentent des blocs a priori non nuls. En tout cas il est vu sous cette forme que $(X - z\mathbb{1})^k$ divise χ_u. □

9.9 Diagonalisation et trigonalisation

Ici encore $\mathbb{K}$ est un corps commutatif.

9.9.1 Matrices semblables

Définition 9.134 (matrices semblables)**.**
Sur l'ensemble $\mathbb{M}_n(\mathbb{K})$ des matrices $n \times n$ à coefficients dans $\mathbb{K}$ nous introduisons la relation d'équivalence $A \sim B$ si et seulement s'il existe une matrice $P \in \mathrm{GL}(n, \mathbb{K})$ telle que $B = P^{-1}AP$. Deux matrices équivalentes en ce sens sont dites ***semblables****.*

Le polynôme caractéristique est un invariant sous les similitudes. En effet si P est une matrice inversible,

$$\begin{aligned} \chi_{PAP^{-1}} &= \det(PAP^{-1} - \lambda X) && \text{(9.298a)} \\ &= \det\left(P^{-1}(PAP^{-1} - \lambda X)P^{-1}\right) && \text{(9.298b)} \\ &= \det(A - \lambda X). && \text{(9.298c)} \end{aligned}$$

La permutation de lignes ou de colonnes ne sont pas de similitudes, comme le montrent les exemples suivants :

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}. \tag{9.299}$$

Nous avons $\chi_A = x^2 - 5x - 2$ tandis que $\chi_B = x^2 - 5x + 2$ alors que le polynôme caractéristique est un invariant de similitude.

9.9.2 Endomorphismes nilpotents

La **trace** d'une matrice $A \in \mathbb{M}(n, \mathbb{K})$ est la somme de ses éléments diagonaux :

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} A_{ii}. \tag{9.300}$$

Une propriété importante est son invariance cyclique.

Lemme 9.135.
Si A et B sont des matrices carré, alors $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.
La trace est un invariant de similitude.

Démonstration. C'est un simple calcul :

$$\mathrm{Tr}(AB) = \sum_{ik} A_{ik}B_{ki} = \sum_{ik} A_{ki}B_{ik} = \sum_{ik} B_{ik}A_{ki} = \sum_{i}(BA)_{ii} = \mathrm{Tr}(BA) \tag{9.301}$$

où nous avons simplement renommé les indices $i \leftrightarrow k$.

En particulier, la trace est un invariant de similitude parce que $\mathrm{Tr}(ABA^{-1}) = \mathrm{Tr}(A^{-1}AB) = \mathrm{Tr}(B)$. □

La trace étant un invariant de similitude, nous pouvons donc définir la **trace** comme étant la trace de sa matrice dans une base quelconque. Si la matrice est diagonalisable, alors la trace est la somme des valeurs propres.

Lemme 9.136 ([34])**.**
L'endomorphisme $u \in \mathrm{End}(\mathbb{C}^n)$ est nilpotent si et seulement si $\mathrm{Tr}(u^p) = 0$ *pour tout p.*

Démonstration. Supposons que u est nilpotent. Alors ses valeurs propres sont toutes nulles et celles de u^p le sont également. La trace étant la somme des valeurs propres, nous avons alors tout de suite $\mathrm{Tr}(u^p) = 0$.

Supposons maintenant que $\mathrm{Tr}(u^p) = 0$ pour tout p. Le polynôme caractéristique (9.276) est

$$\chi_u = (-1)^n X^\alpha (X - \lambda_1)^{\alpha_1} \dots (X - \alpha_r)^{\alpha_r}. \tag{9.302}$$

où les λ_i $(i = 1, \dots, r)$ sont les valeurs propres non nulles distinctes de u.

Il est vite vu que le coefficient de X^{n-1} dans χ_u est $-\mathrm{Tr}(u)$ parce que le coefficient de X^{n-1} se calcule en prenant tous les X sauf une fois $-\lambda_i$. D'autre part le polynôme caractéristique de u^p est le même que celui de u, en remplaçant λ_i par λ_i^p ; cela est dû au fait que si v est vecteur propre de valeur propre λ, alors $u^p v = \lambda^p v$.

Par l'équation (9.302), nous voyons que le coefficient du terme X^{n-1} dans les polynôme caractéristique est

$$0 = \mathrm{Tr}(u^p) = \alpha_1\lambda_1^p + \dots + \alpha_r\lambda_r^p. \tag{9.303}$$

Donc les nombres $(\alpha_1, \dots, \alpha_r)$ est une solution non triviale [25] du système

$$\begin{cases} \alpha_1 X_1 + \dots + \lambda_r X_r = 0 & \text{(9.304a)} \\ \quad\vdots & \text{(9.304b)} \\ \lambda_1^r X_1 + \dots + \lambda_r^r X_r = 0. & \text{(9.304c)} \end{cases}$$

Cela sont les équations (9.303) écrites avec $p = 1, \dots, r$. Le déterminant de ce système est

$$\lambda_1 \dots \lambda_r \det \begin{pmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_1 \\ \vdots & & \vdots \\ \lambda_1^{r-1} & \dots & \lambda_r^{r-1} \end{pmatrix} \neq 0, \tag{9.305}$$

25. Si $\alpha_1 = \dots = \alpha_r = 0$, alors les valeurs propres sont toutes nulles et la matrice est en réalité nulle dès le départ.

qui est un déterminant de Vandermonde (proposition 9.12) valant

$$0 = \lambda_1 \dots \lambda_r \prod_{1 \leqslant i \leqslant j \leqslant r} (\lambda_i - \lambda_j). \tag{9.306}$$

Étant donné que les λ_i sont distincts et non nuls, nous avons une contradiction et nous devons conclure que $(\alpha_1, \dots, \alpha_r)$ était une solution triviale du système (9.304). □

Proposition 9.137 ([101])**.**
Soit un $\mathbb{K}$-espace vectoriel E. Un endomorphisme $u \in \mathrm{End}(E)$ est nilpotent si et seulement s'il existe une base de E dans laquelle la matrice de u est strictement triangulaire supérieure.

Démonstration. $\Rightarrow$ Nous faisons la démonstration par récurrence sur la dimension de E. Lorsque $n = 1$ nous avons $u = (a)$ avec $a \in \mathbb{K}$. Vu que $a^k = 0$ pour un certain k nous avons $a = 0$ parce qu'un corps est toujours un anneau intègre [26].

Lorsque $\dim(E) = n$ nous savons que u a un noyau non réduit au vecteur nul (parce qu'il est nilpotent). Soit donc un vecteur non nul $x \in \ker(u)$ et une base

$$\{x, e_2, \dots, e_n\} \tag{9.307}$$

donnée par le théorème de la base incomplète 7.10. La matrice de u dans cette base s'écrit

$$\left(\begin{array}{c|ccc} 0 & * & * & * \\ \hline 0 & & & \\ 0 & & A & \\ 0 & & & \end{array}\right). \tag{9.308}$$

Un tout petit peu de calcul de produit de matrice montre que la matrice de u^k est de la forme

$$\left(\begin{array}{c|ccc} 0 & * & * & * \\ \hline 0 & & & \\ 0 & & A^k & \\ 0 & & & \end{array}\right). \tag{9.309}$$

Étant donné que u est nilpotente, la matrice A l'est aussi. L'hypothèse de récurrence dit alors que A est strictement triangulaire supérieure (ou en tout cas peut le devenir par un changement de base adéquat).

$\Leftarrow$ Lorsqu'une matrice est triangulaire supérieure stricte, elle applique

$$\mathrm{Span}\{e_1, \dots, e_k\} \to \mathrm{Span}\{e_1, \dots, e_{k-1}\}. \tag{9.310}$$

Donc tout vecteur finit sur zéro si on lui applique u assez souvent.

□

Proposition 9.138 (Thème 40)**.**
Soit E un espace de Banach (espace vectoriel normé complet). Si $A \in \mathcal{L}(E, E)$ est nilpotente, alors $(\mathbb{1} - A)$ est inversible et son inverse est donné par

$$(\mathbb{1} - A)^{-1} = \sum_{k=0}^{\infty} A^k, \tag{9.311}$$

où l'infini peut évidemment être remplacé par l'ordre de nilpotence de A.

26. Lemme 4.38.

Démonstration. En ce qui concerne la convergence de la somme, elle ne fait pas de doutes parce que A étant nilpotente, la somme contient seulement une quantité finie de termes non nuls.

Montrons à présent que la somme est l'inverse de $\mathbb{1} - A$ en multipliant terme à terme :

$$\sum_{k=0}^{n} A^k(\mathbb{1} - A) = \sum_{k=0}^{n}(A^k - A^{k+1}) = \mathbb{1} - A^{n+1}. \tag{9.312}$$

Par conséquent

$$\|\mathbb{1} - \sum_{k=0}^{n} A^k(\mathbb{1} - A)\| = \|A^{n+1}\| \to 0. \tag{9.313}$$

La dernière limite est en réalité une égalité pour n assez grand. □

Proposition 9.139.
Soit $A \in \mathrm{GL}(n, \mathbb{C})$. La suite $(A^k)_{k\in\mathbb{Z}}$ est bornée si et seulement si A est diagonalisable et $\mathrm{Spec}(A) \subset \mathrm{S}^1$.

Démonstration. Si A est diagonalisable avec les valeurs propres λ_i de norme 1 dans $\mathbb{C}$, alors A^k est la matrice diagonale avec les λ_i^k sur la diagonale. Cela reste borné pour toute valeur entière de k.

En ce qui concerne l'autre sens, nous supposons encore que

$$A = \begin{pmatrix} \lambda_1\mathbb{1} + N_1 & & \\ & \ddots & \\ & & \lambda_s\mathbb{1} + N_s \end{pmatrix}, \tag{9.314}$$

et nous regardons un des blocs. Nous voulons prouver que $N = 0$ et que $|\lambda| = 1$.

Nous commençons par regarder ce qu'implique le fait que $(\lambda\mathbb{1} + N)^n$ reste borné pour $n > 0$. En notant r l'ordre de nilpotence de N, nous avons le développement

$$(\lambda\mathbb{1} + N)^n = \sum_{k=0}^{r-1} \binom{n}{k} N^k\lambda^{n-k}. \tag{9.315}$$

Par la proposition 9.137, une matrice nilpotente s'écrit dans une base sous la forme

$$N = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \tag{9.316}$$

et effectuer A^k revient à décaler la diagonale de 1. Donc la famille

$$\{\mathbb{1}, N, \dots, N^{r-1}\} \tag{9.317}$$

est libre. Par conséquent la suite $(\lambda\mathbb{1} + N)^n$ restera bornée si et seulement si chacun des termes

$$\binom{n}{k} N^k\lambda^{n-k} \tag{9.318}$$

reste borné. Le premier terme étant $\lambda^n\mathbb{1}$, nous avons obligatoirement $|\lambda| \leqslant 1$. Si $|\lambda| < 1$, alors le coefficient $\binom{n}{k}\lambda^{n-k}$ tend vers zéro. Si $|\lambda| = 1$ par contre ce coefficient tend vers l'infini et la seule façon pour que (9.318) reste borné est que $N = 0$. Nous avons donc deux possibilités :

— $|\lambda| < 1$
— $|\lambda| = 1$ et $N = 0$.

Nous nous tournons maintenant sur la contrainte que $(\lambda\mathbb{1}+N)^n$ doive rester borné pour $n<0$. Nous avons

$$\lambda\mathbb{1}+N=\lambda(\mathbb{1}+\lambda^{-1}N), \tag{9.319}$$

et nous pouvons appliquer la proposition 9.138 à l'opérateur nilpotent $-\lambda^{-1}N$ pour avoir

$$(\mathbb{1}+\lambda^{-1}N)^{-1}=\mathbb{1}+\sum_{k=1}^{\infty}(-\lambda)^{-1}N^k. \tag{9.320}$$

Ceci pour dire que $(\lambda\mathbb{1}+N)^{-1}=\lambda^{-1}(\mathbb{1}+\lambda^{-1}N')$ pour une autre matrice nilpotente N'. Le travail déjà fait, appliqué à λ^{-1} et N', nous donne deux possibilités :

— $|\lambda^{-1}|<1$
— $|\lambda^{-1}|=1$ et $N'=0$.

La possibilité $|\lambda^{-1}|<1$ est exclue parce qu'elle impliquerait $|\lambda|>1$ qui avait déjà été exclu. Il ne reste donc que la possibilité $|\lambda|=1$ et $N=N'=0$. □

9.9.3 Endomorphismes diagonalisables

Définition 9.140.
*Une matrice est **diagonalisable** si elle est semblable à une matrice diagonale.*

Lemme 9.141.
Une matrice triangulaire supérieure avec des 1 *sur la diagonale n'est diagonalisable que si elle est diagonale (c'est à dire si elle est la matrice unité).*

Démonstration. Si A est une matrice triangulaire supérieure de taille n telle que $A_{ii}=1$, alors $\det(A-\lambda\mathbb{1})=(1-\lambda)^n$, ce qui signifie que $\mathrm{Spec}(A)=\{1\}$. Pour la diagonaliser, il faudrait une matrice $P\in\mathrm{GL}(n,\mathbb{K})$ telle que $\mathbb{1}=P^{-1}AP$, ce qui est uniquement possible si $A=\mathbb{1}$. □

Lemme 9.142.
Soit F un sous-espace stable par u. Soit une décomposition du polynôme minimal

$$\mu_u=P_1^{n_1}\dots P_r^{n_r} \tag{9.321}$$

où les P_i sont des polynômes irréductibles unitaires distincts. Si nous posons $E_i=\ker P_i^{n_i}$, alors

$$F=(F\cap E_1)\oplus\dots\oplus(F\cap E_r). \tag{9.322}$$

Théorème 9.143.
Soit E, un espace vectoriel de dimension n sur le corps commutatif $\mathbb{K}$ et $u\in\mathrm{End}(E)$. Les propriétés suivantes sont équivalentes.

(1) L'endomorphisme u est diagonalisable.
(2) Il existe un polynôme $P\in\mathbb{K}[X]$ non constant, scindé sur $\mathbb{K}$ dont toutes les racines sont simples tel que $P(u)=0$.
(3) Le polynôme minimal μ_u est scindé sur $\mathbb{K}$ et toutes ses racines sont simples[27].
(4) Tout sous-espace de E possède un supplémentaire stable par u.
(5) Dans une base adaptée, la matrice de u est diagonale et les éléments diagonaux sont ses valeurs propres.

Démonstration. Plein d'implications à prouver.

27. Le polynôme *caractéristique*, lui, n'a pas spécialement ses racines simples ; il peut encore être de la forme

$$\chi_u(X)=\prod_{i=1}^{r}(X-\lambda_i)^{\alpha_i}, \tag{9.323}$$

mais alors $\dim(E_{\lambda_i})=\alpha_i$.

(2) implique (3) Étant donné que $P(u) = 0$, il est dans l'idéal des polynôme annulateurs de u, et le polynôme minimal μ_u le divise parce que l'idéal des polynôme annulateurs est généré par μ_u par le théorème 5.37.

(3) implique (1) Étant donné que le polynôme minimal est scindé à racines simples, il s'écrit sous forme de produits de monômes tous distincts, c'est à dire

$$\mu_u(X) = (X - \lambda_1)\dots(X - \lambda_r) \tag{9.324}$$

où les λ_i sont des éléments distincts de $\mathbb{K}$. Étant donné que $\mu_u(u) = 0$, le théorème de décomposition des noyaux (théorème 9.102) nous enseigne que

$$E = \ker(u - \lambda_1) \oplus \dots \oplus \ker(u - \lambda_r). \tag{9.325}$$

Mais $\ker(u - \lambda_i)$ est l'espace propre $E_{\lambda_i}(u)$. Donc u est diagonalisable.

(1) implique (4) Soit $\{e_1, \dots, e_n\}$ une base qui diagonalise u, soit F un sous-espace de E un $\{f_1, \dots, f_r\}$ une base de F. Par le théorème 7.14 (qui généralise le théorème de la base incomplète), nous pouvons compléter la base de F par des éléments de la base $\{e_i\}$. Le complément ainsi construit est invariant par u.

(4) implique (1) En dimension un, tout endomorphisme est diagonalisable, nous supposons donc que $\dim E = n \geqslant 2$. Nous procédons par récurrence sur le nombre de vecteurs propres connus de u. Supposons avoir déjà trouvé p vecteurs propres $e_1, \dots, e_p$ de u. Considérons H, un hyperplan qui contient les vecteurs $e_1, \dots, e_p$. Soit F un supplémentaire de H stable par u ; par construction $\dim F = 1$ et si $e_{p+1} \in F$, il doit être vecteur propre de u.

(1) implique (2) Nous supposons maintenant que u est diagonalisable. Soient $\lambda_1, \dots, \lambda_r$ les valeurs propres deux à deux distinctes, et considérons le polynôme

$$P(x) = (X - \lambda_1)\dots(X - \lambda_r). \tag{9.326}$$

Alors $P(u) = 0$. En effet si e_i est un vecteur propre pour la valeur propre λ_i,

$$P(u)e_i = \prod_{j \neq i}(u - \lambda_j) \circ (u - \lambda_i)e_i = 0 \tag{9.327}$$

par le lemme 9.101. Par conséquent $P(u)$ s'annule sur une base.

(5) implique (2) Si la matrice A est diagonale alors le polynôme $P = \prod_{i=1}^{n}(A - A_{ii}\mathbb{1})$ est annulateur de A.

(3) implique (5) le polynôme minimal de u s'écrit

$$\mu = (X - \lambda_1)\dots(X - \lambda_r), \tag{9.328}$$

et les espaces E_i du lemme 9.142 sont les espaces propres $E_i = \ker(u - \lambda_i)$. Nous avons donc une somme directe

$$E = E_1 \oplus \dots \oplus E_r. \tag{9.329}$$

Dans chacun des espaces propres, u a une matrice diagonale avec la valeur propre correspondante sur la diagonale. Une base de E constituée d'une base de chacun des espaces propres est donc une base comme nous en cherchons.

□

Corollaire 9.144.
Si u est diagonalisable et si F est une sous-espace stable par u, alors

$$F = \bigoplus_{\lambda} E_\lambda(u) \cap F \tag{9.330}$$

où $E_\lambda(u)$ est l'espace propre de u pour la valeur propre λ. En particulier la restriction de u à F, $u|_F$ est diagonalisable.

Démonstration. Par le théorème 9.143, le polynôme μ_u est scindé et ne possède que des racines simples. Notons le

$$\mu_u(X) = (X - \lambda_1)\dots(X - \lambda_r). \tag{9.331}$$

Les espaces E_i du lemme 9.142 sont maintenant les espaces propres.

En ce qui concerne la diagonalisabilité de $u|_F$, notons que nous avons une base de F composée de vecteurs dans les espaces $E_\lambda(u)$. Cette base de F est une base de vecteurs propres de u. □

Lemme 9.145.
Soit E un $\mathbb{K}$-espace vectoriel et $u \in \mathrm{End}(E)$. Si $\mathrm{Card}\big(\mathrm{Spec}(u)\big) = \dim(E)$ alors u est diagonalisable.

Démonstration. Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres distinctes de u. Nous savons que les espaces propres correspondants sont en somme directe (lemme 9.99). Par conséquent $\mathrm{Span}\{E_{\lambda_i}(u)\}$ est de dimension n est u est diagonalisable. □

Voici un résultat de diagonalisation simultanée. Nous donnerons un résultat de trigonalisation simultanée dans le lemme 9.240.

Proposition 9.146 (Diagonalisation simultanée)**.**
Soit $(u_i)_{i\in I}$ une famille d'endomorphismes qui commutent deux à deux.

(1) Si $i, j \in I$ alors tout sous-espace propre de u_i est stable par u_j. Autrement dit $u_j\big(E_\lambda(u)\big) \subset E_\lambda(u)$.

(2) Si les u_i sont diagonalisables, alors ils le sont simultanément.

Démonstration. Supposons que u_i et u_j commutent et soit x un vecteur propre de u_i : $u_i x = \lambda x$. Nous montrons que $u_j x \in E_\lambda(u)$. Nous avons

$$u_i\big(u_j(x)\big) = u_j\big(u_i(x)\big) = \lambda u_j(x). \tag{9.332}$$

Par conséquent $u_j(x)$ est vecteur propre de u_i de valeur propre λ.

Montrons maintenant l'affirmation à propos des endomorphismes simultanément diagonalisables. Si $\dim E = 1$, le résultat est évident. Nous supposons également qu'aucun des u_i n'est multiple de l'identité. Nous effectuons une récurrence sur la dimension.

Soit u_0 un des u_i et considérons ses valeurs propres deux à deux distinctes $\lambda_1, \dots, \lambda_r$. Pour chaque k nous avons

$$E_{\lambda_k}(u_0) \neq E, \tag{9.333}$$

sinon u_0 serait un multiple de l'identité. Par contre le fait que u_0 soit diagonalisable permet de décomposer E en espaces propres de u_0 :

$$E = \bigoplus_k E_{\lambda_k}(u_0). \tag{9.334}$$

Ce que nous allons faire est de simultanément diagonaliser les $(u_i)_{i\in I}$ sur chacun des E_{λ_k} séparément. Par le point (1), nous avons $u_i\colon E_{\lambda_k}(u_0) \to E_{\lambda_k}(u_0)$, et nous pouvons considérer la famille d'opérateurs

$$\Big(u_i|_{E_{\lambda_k}(u_0)}\Big)_{i\in I}. \tag{9.335}$$

Ce sont tous des opérateurs qui commutent et qui agissent sur un espace de dimension plus petite. Par hypothèse de récurrence nous avons une base de $E_{\lambda_k}(u_0)$ qui diagonalise tous les u_i. □

Exemple 9.147
Soit un espace vectoriel sur un corps $\mathbb{K}$. Un opérateur **involutif** est un opérateur différent de l'identité dont le carré est l'identité. Typiquement une symétrie orthogonale dans $\mathbb{R}^3$. Le polynôme caractéristique d'une involution est $X^2 - 1 = (X + 1)(X - 1)$.

Tant que $1 \neq -1$, $X^1 - 1$ est donc scindé à racines simples et les involutions sont diagonalisables (9.143). Cependant si le corps est de caractéristique 2, alors $X^2 - 1 = (X+1)^2$ et l'involution n'est plus diagonalisable.

Par exemple si le corps est de caractéristique 2, nous avons

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{9.336a}$$

$$A^1 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{9.336b}$$

Ce A est donc une involution mais n'est pas diagonalisable. △

9.9.4 Diagonalisation : cas complexe, pas toujours

Il n'est pas vrai qu'une matrice de $\mathbb{M}(n, \mathbb{C})$ soit toujours diagonalisable. En effet le théorème 9.143(3) dit qu'une matrice est diagonalisable si et seulement si son polynôme minimal est scindé à racines simples. Certes sur $\mathbb{C}$ le polynôme minimal sera scindé, mais il ne sera pas spécialement à racines simples.

Exemple 9.148
La matrice

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{9.337}$$

a pour polynôme caractéristique $\chi_A(X) = X^2$. Cela est également son polynôme minimal, et ce n'est pas à racine simple.

Il est par ailleurs facile de voir que le seul espace propre de A est $\text{Span}\{(1,0)\}$ (ici le span est sur $\mathbb{C}$). Donc l'espace $\mathbb{C}^2$ ne possède pas de base de vecteurs propres de A. △

Ce qui est vrai, c'est que le polynôme caractéristique a des racines, et que ces racines correspondent à des vecteurs propres. Mais il n'y a pas toujours autant de vecteurs propres que la multiplicité des racines.

9.9.5 Trigonalisation : généralités

Définition 9.149 ([102]).
*Une matrice dans $\mathbb{M}(n, \mathbb{K})$ est **trigonalisable** lorsqu'elle est semblable*[28] *à une matrice triangulaire supérieure.*

Proposition 9.150 (Trigonalisation et polynôme caractéristique scindé).
Soit u un endomorphisme d'un espace vectoriel E sur le corps $\mathbb{K}$. Les faits suivants sont équivalents.

(1) L'endomorphisme u est trigonalisable (auquel cas les valeurs propres sont sur la diagonale).

(2) Le polynôme caractéristique de u est scindé[29].

Démonstration. **(2)⇒(1)** Nous avons par hypothèse que

$$\chi_u(X) = \prod_{i=1}^{r} (X - \lambda_i)^{\alpha_i} \tag{9.338}$$

où les λ_i sont les valeurs propres de u. Le théorème de Cayley-Hamilton 9.130 dit que $\chi_u(u) = 0$, ce qui permet d'utiliser le théorème de décomposition des noyaux 9.102 :

$$E = \ker(X - \lambda_1)^{\alpha_1} \oplus \ldots \oplus \ker(X - \lambda_r)^{\alpha_r}. \tag{9.339}$$

28. Définition 9.134.
29. Définition 5.33.

Les espaces $F_{\lambda_i}(u) = \ker(X - \lambda_i)^{\alpha_i}$ sont les espaces caractéristiques de u, ce qui fait que $u-\lambda_i\mathbb{1}$ est nilpotent sur $F_{\lambda_i}(u)$. L'endomorphisme $u-\lambda_i\mathbb{1}$ est donc strictement trigonalisable supérieur sur son bloc[30]. Cela signifie que u est triangulaire supérieure avec les valeurs propres sur la diagonale.

(1)⇒(2) C'est immédiat parce que le déterminant d'une matrice triangulaire est le produit des éléments de sa diagonale.

□

Remarque 9.151.
La méthode des pivots de Gauss[31] certes permet de trigonaliser n'importe quoi, mais elle ne correspond pas à un changement de base. Autrement dit, les pivots de Gauss ne sont pas de similitudes.

C'est là qu'il faut bien avoir en tête la différence entre *équivalence* et *similarité* (définition 7.27). Lorsqu'on parle de changement de base, de matrice trigonalisable ou diagonalisable, nous parlons de similarité et non d'équivalence.

9.9.6 Trigonalisation : cas complexe

La proposition 9.150 dit déjà que tous les endomorphismes sont trigonalisables sur $\mathbb{C}$. Nous allons aller plus loin et montrer que la trigonalisation peut être effectuée à l'aide d'une matrice unitaire.

Une démonstration alternative passant par le polynôme caractéristique sera présentée dans la remarque 9.155 utilisant la proposition 9.150.

Lemme 9.152 (Lemme de Schur complexe, trigonisation[103])**.**
Si $A \in \mathbb{M}(n, \mathbb{C})$, il existe une matrice unitaire U telle que UAU^{-1} soit triangulaire supérieure.

Démonstration. Étant donné que $\mathbb{C}$ est algébriquement clos, nous pouvons toujours considérer un vecteur propre v_1 de A, de valeur propre λ_1. Nous pouvons utiliser un procédé de Gram-Schmidt pour construire une base orthonormée $\{v, u_2, \dots, u_n\}$ de $\mathbb{R}^n$, et la matrice (unitaire)

$$Q = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ v & u_2 & \cdots & u_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}. \tag{9.340}$$

Nous avons $Q^{-1}AQe_1 = Q^{-1}Av = \lambda Q^{-1}v = \lambda e_1$, par conséquent la matrice $Q^{-1}AQ$ est de la forme

$$Q^{-1}AQ = \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix} \tag{9.341}$$

où $*$ représente une ligne quelconque et A_1 est une matrice de $\mathbb{M}(n-1, \mathbb{C})$. Nous pouvons donc répéter le processus sur A_1 et obtenir une matrice triangulaire supérieure (nous utilisons le fait qu'un produit de matrices orthogonales est une matrice orthogonale). □

En particulier les matrices hermitiennes, anti-hermitiennes et unitaires sont trigonalisables par une matrice unitaire, qui peut être choisie de déterminant 1.

Lemme 9.153.
Soit $A \in \mathbb{M}(n, \mathbb{C})$ et une matrice unitaire U telle que $A = UTU^{-1}$ où T est triangulaire.

(1) En ce qui concerne les polynômes caractéristiques, $\chi_A = \chi_T$.

(2) Pour les spectres, $\mathrm{Spec}(A) = \mathrm{Spec}(T)$.

(3) Les valeurs propres de A sont les éléments diagonaux de T.

30. Proposition 9.137.
31. Le lemme 7.28.

Démonstration. Vu que U commute évidemment avec $\mathbb{1}$ nous avons

$$\chi_A(\lambda) = \det(A - \lambda\mathbb{1}) = \det(UTU^{-1} - \lambda\mathbb{1}) = \det\big(U(T - \lambda\mathbb{1})U^{-1}\big). \tag{9.342}$$

À ce niveau nous utilisons le fait que le déterminant soit multiplicatif 9.10 pour conclure :

$$\chi_A(\lambda) = \det\big(U(T - \lambda\mathbb{1})U^{-1}\big) = \det(U)\det(T - \lambda\mathbb{1})\det(U^{-1}) = \det(T - \lambda\mathbb{1}) = \chi_T(\lambda). \tag{9.343}$$

Pour les spectres, l'égalité des polynômes caractéristique implique l'égalité des spectres parce que les valeurs propres sont les racines du polynôme caractéristique par le théorème 9.125.

Les valeurs propres d'une matrice triangulaire sont les valeurs sur la diagonale. □

Remarque 9.154.
Le lemme mentionne le fait que les valeurs propres de A sont les éléments diagonaux de T. Mais attention : ceci ne dit rien au niveau des multiplicités géométriques. Un nombre peut être cinq fois sur la diagonale de T alors que l'espace propre correspondant pour A n'est que de dimension 1. Exemple : la matrice

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{9.344}$$

a deux 1 sur la diagonale. Le nombre 1 est bien une valeur propre de A, mais le système

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{9.345}$$

donne $y = 0$ et donc un espace propre de dimension seulement 1.

Remarque 9.155.
Si $\mathbb{K}$ est algébriquement clos (comme $\mathbb{C}$ par exemple), alors tous les polynômes sont scindés et toutes les matrices sont trigonalisables[32]. Un exemple un peu simple de cela est la matrice

$$u = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{9.346}$$

Le polynôme caractéristique est $\chi_u(X) = X^2 + 1$ et les valeurs propres sont $\pm i$. Il est vite vu que dans la base

$$\{\begin{pmatrix} i \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ i \end{pmatrix}\} \tag{9.347}$$

de $\mathbb{C}^2$, la matrice u se note $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$.

Remarque 9.156.
Cela nous donne une autre façon de prouver qu'une matrice nilpotente de $\mathbb{M}(n, \mathbb{C})$ ou $\mathbb{M}(n, \mathbb{R})$ est trigonalisable[104]. D'abord dans $\mathbb{M}(n, \mathbb{C})$, toutes les matrices sont trigonalisables[33], et les valeurs propres arrivent sur la diagonale. Mais comme les valeurs propres d'une matrice nilpotente sont zéro, elle est triangulaire stricte. Par ailleurs son polynôme caractéristique est alors X^n.

Ensuite si $u \in \mathbb{M}(n, \mathbb{R})$ nous pouvons voir u comme une matrice dans $\mathbb{M}(n, \mathbb{C})$ et y calculer son polynôme caractéristique qui sera tout de même X^n. Ce polynôme étant scindé, la proposition 9.150 nous assure que u est trigonalisable. Une fois de plus, les valeurs propres étant sur la diagonale, elle est triangulaire supérieure stricte.

Corollaire 9.157.
Le polynôme caractéristique[34] *sur $\mathbb{C}$ d'une matrice s'écrit sous la forme*

$$\chi_A(X) = \prod_{i=1}^{r}(X - \lambda_i)^{m_i} \tag{9.348}$$

où les λ_i sont les valeurs propres distinctes de A et m_i sont les multiplicités correspondantes.

32. La proposition 9.150 montre cela, et le lemme de Schur complexe 9.152 va un peu plus loin et précise que la trigonalisation peut être faite par une matrice unitaire.

33. Parce que le polynôme caractéristique est scindé, voir la proposition 9.150..

34. Définition 9.121.

Démonstration. Le lemme 9.152 nous donne l'existence d'une base de trigonalisation ; dans cette base les valeurs propres de A sont sur la diagonale et nous avons

$$\chi_A(X) = \det(A - X\mathbb{1}) = \det \begin{pmatrix} X - \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & X - \lambda_r \end{pmatrix}, \tag{9.349}$$

qui vaut bien le produit annoncé. □

Corollaire 9.158.
Si $A \in \mathbb{M}(n, \mathbb{C})$ et $k \in \mathbb{N}$ alors

$$\operatorname{Spec}(A^k) = \{\lambda^k \text{ tel que } \lambda \in \operatorname{Spec}(A)\}. \tag{9.350}$$

Démonstration. Par le lemme 9.152 nous avons une matrice unitaire U et une triangulaire T telles que $A = UTU^{-1}$. En passant à a puissance k nous avons aussi

$$A^k = UT^kU^{-1}. \tag{9.351}$$

Donc le spectre de A^k est celui de T^k (lemme 9.153 et le fait qu'une puissance d'une matrice triangulaire est encore triangulaire). Or les éléments diagonaux de T^k sont les puissance k^{e}des éléments diagonaux de T, qui sont les valeurs propres de A. □

9.9.7 Diagonalisation : cas complexe, ce qu'on a

Lemme 9.159 (Théorème spectral hermitien)**.**
Pour un opérateur hermitien[35]*,*

(1) le spectre est réel,

(2) deux vecteurs propres à des valeurs propres distinctes sont orthogonales[36]*.*

Démonstration. Soit v un vecteur de valeur propre λ. Nous avons d'une part

$$\langle Av, v\rangle = \lambda\langle v, v\rangle = \lambda\|v\|^2, \tag{9.352}$$

et d'autre part, en utilisant le fait que A est hermitien,

$$\langle Av, v\rangle = \langle v, A^*v\rangle = \langle v, Av\rangle = \bar{\lambda}\|v\|^2, \tag{9.353}$$

par conséquent $\lambda = \bar{\lambda}$ parce que $v \neq 0$.

Soient λ_i et v_i $(i = 1, 2)$ deux valeurs propres de A avec leurs vecteurs propres correspondants. Alors d'une part

$$\langle Av_1, v_2\rangle = \lambda_1\langle v_1, v_2\rangle, \tag{9.354}$$

et d'autre part

$$\langle Av_1, v_2\rangle = \langle v_1, Av_2\rangle = \lambda_2\langle v_1, v_2\rangle. \tag{9.355}$$

Nous avons utilisé le fait que λ_2 était réel. Par conséquent, soit $\lambda_1 = \lambda_2$, soit $\langle v_1, v_2\rangle = 0$. □

Remarque 9.160.
Un opérateur de la forme A^*A est évidemment hermitien. De plus ses valeurs propres sont toutes positives parce que si $A^*Ax = \lambda v$ alors

$$0 \leqslant \langle Av, Av\rangle = \langle A^*Av, v\rangle = \lambda\langle v, v\rangle. \tag{9.356}$$

Donc $\lambda \geqslant 0$.

35. Définition 9.45.
36. Pour la forme (9.130).

Définition 9.161.
Un endomorphisme est ***normal*** *s'il commute avec son adjoint.*

Théorème 9.162 (Théorème spectral pour les matrices normales[37][105, 106])**.**
Soit $A \in \mathbb{M}(n, \mathbb{C})$ une matrice de valeurs propres $\lambda_1, \dots, \lambda_n$ (non spécialement distinctes). Alors les conditions suivantes sont équivalentes :

(1) A est normale,
(2) A se diagonalise par une matrice unitaire,
(3) $\sum_{i,j=1}^{n} |A_{ij}|^2 = \sum_{j=1}^{n} |\lambda_j|^2$,
(4) il existe une base orthonormale de vecteurs propres de A.

Démonstration. Nous allons nous contenter de prouver (1)⇔(2).

Soit Q la matrice unitaire donnée par la décomposition de Schur (lemme 9.152) : $A = QTQ^{-1}$. Étant donné que A est normale nous avons

$$QTT^*Q^{-1} = QT^*TQ^{-1}, \tag{9.357}$$

ce qui montre que T est également normale. Or une matrice triangulaire supérieure normale est diagonale. En effet nous avons $T_{ij} = 0$ lorsque $i > j$ et

$$(TT^*)_{ii} = (T^*T)_{ii} = \sum_{k=1}^{n} |T_{ki}|^2 = \sum_{k=1}^{n} |T_{ik}|^2. \tag{9.358}$$

Écrivons cela pour $i = 1$ en tenant compte de $|T_{k1}|^2 = 0$ pour $k = 2, \dots, n$,

$$|T_{11}|^2 = |T_{11}|^2 + |T_{12}|^2 + \cdots + |T_{1n}|^2, \tag{9.359}$$

ce qui implique que T_{11} est le seul non nul parmi les T_{1k}. En continuant de la sorte avec $i = 2, \dots, n$ nous trouvons que T est diagonale.

Dans l'autre sens, si A se diagonalise par une matrice unitaire, $UAU^* = D$, nous avons

$$DD^* = UAA^*U^* \tag{9.360}$$

et

$$D^*D = UA^*AU^*, \tag{9.361}$$

qui ce prouve que A est normale. □

Tant que nous en sommes à parler de spectre de matrices hermitiennes... Soit une matrice inversible $A \in \mathrm{GL}(n, \mathbb{C})$. La matrice A^*A est hermitienne[38] et le théorème 9.159 nous assure que ses valeurs propres sont réelles. Par la remarque 9.160, ses valeurs propres sont même positives.

Lemme 9.163 ([107])**.**
Si A est une matrice carré et inversible,

$$\mathrm{Spec}(A^*A) = \mathrm{Spec}(AA^*) \tag{9.362}$$

Démonstration. Nous allons montrer l'égalité des polynômes caractéristiques. D'abord une simple multiplication montre que

$$(A^*A - \lambda\mathbb{1})A^{-1} = A^{-1}(AA^* - \lambda\mathbb{1}). \tag{9.363}$$

Nous prenons le déterminant de cette égalité en utilisant les propriétés 9.10(1) et (3) :

$$\det(A^*A - \lambda\mathbb{1})\det(A^{-1}) = \det(A^{-1})\det(AA^* - \lambda\mathbb{1}). \tag{9.364}$$

En simplifiant par $\det(A^{-1})$ (qui est non nul parce que A est inversible) nous obtenons l'égalité des polynômes caractéristiques et donc l'égalité des spectres. □

37. Définition 9.161
38. Définition 9.45.

9.9.8 Diagonalisation : cas réel

Lemme 9.164 (Lemme de Schur réel)**.**
Soit $A \in \mathbb{M}(n,\mathbb{R})$. Il existe une matrice orthogonale Q telle que $Q^{-1}AQ$ soit de la forme

$$QAQ^{-1} = \begin{pmatrix} \lambda_1 & * & * & * & * \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \lambda_r & * & * \\ 0 & 0 & 0 & \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} & * \\ 0 & 0 & 0 & 0 & \begin{pmatrix} a_s & b_s \\ c_s & d_s \end{pmatrix} \end{pmatrix}. \tag{9.365}$$

Le déterminant de A est le produit des déterminants des blocs diagonaux et les valeurs propres de A sont les $\lambda_1,\dots,\lambda_r$ et celles de ces blocs.

Démonstration. Si la matrice A a des valeurs propres réelles, nous procédons comme dans le cas complexe. Cela nous fournit le partie véritablement triangulaire avec les valeurs propres $\lambda_1,\dots,\lambda_r$ sur la diagonale. Supposons donc que A n'a pas de valeurs propres réelles. Soit donc $\alpha+i\beta$ une valeur propre ($\beta\neq 0$) et $u+iv$ un vecteur propre correspondant où u et v sont des vecteurs réels. Nous avons

$$Au+iAv = A(u+iv) = (\alpha+i\beta)(u+iv) = \alpha u-\beta v+i(\alpha v+\beta v), \tag{9.366}$$

et en égalisant les parties réelles et imaginaires,

$$Au = \alpha u-\beta v \tag{9.367a}$$
$$Av = \alpha v+\beta u. \tag{9.367b}$$

Sur ces relations nous voyons que ni u ni v ne sont nuls. De plus u et v sont linéairement indépendants (sur $\mathbb{R}$), en effet si $v=\lambda u$ nous aurions $Au=\alpha u-\beta\lambda u=(\alpha-\beta\lambda)u$, ce qui serait une valeur propre réelle alors que nous avions supposé avoir déjà épuisé toutes les valeurs propres réelles.

Étant donné que u et v sont deux vecteurs réels non nuls et linéairement indépendants, nous pouvons trouver une base orthonormée $\{q_1,q_2\}$ de $\text{Span}\{u,v\}$. Nous pouvons étendre ces deux vecteurs en une base orthonormée $\{q_1,q_2,q_3,\dots,q_n\}$ de $\mathbb{R}^n$. Nous considérons à présent la matrice orthogonale dont les colonnes sont formées de ces vecteurs : $Q=[q_1\,q_2\,\dots\,q_n]$.

L'espace $\text{Span}\{e_1,e_2\}$ est stable par $Q^{-1}AQ$, en effet nous avons

$$Q^{-1}AQe_1 = Q^{-1}Aq_1 = Q^{-1}(aq_1+bq_2) = ae_1+be_2. \tag{9.368}$$

La matrice $Q^{-1}AQ$ est donc de la forme

$$Q^{-1}AQ = \begin{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} & C_1 \\ 0 & A_1 \end{pmatrix} \tag{9.369}$$

où C_1 est une matrice réelle $2\times(n-1)$ quelconque et A_1 est une matrice réelle $(n-2)\times(n-2)$. Nous pouvons appliquer une récurrence sur la dimension pour poursuivre.

Notons que si A n'a pas de valeurs propres réelles, elle est automatiquement d'ordre pair parce que les valeurs propres complexes viennent par couple complexes conjuguées.

En ce qui concerne les valeurs propres, il est facile de voir en regardant (9.365) que les valeurs propres sont celles des blocs diagonaux. Étant donné que QAQ^{-1} et A ont même polynôme caractéristique, ce sont les valeurs propres de A. □

Théorème 9.165 (Théorème spectral, matrice symétrique[2])**.**
Une matrice symétrique réelle,

(1) a un spectre contenu dans $\mathbb{R}$

(2) est diagonalisable par une matrice orthogonale.

Si M *est une matrice symétrique réelle alors* $\mathbb{R}^n$ *possède une base orthonormée de vecteurs propres de* M*.*

Démonstration. Soit A une matrice réelle symétrique. Si λ est une valeur propre complexe pour le vecteur propre complexe v, alors d'une part $\langle Av, v\rangle = \lambda\langle v, v\rangle$ et d'autre part $\langle Av, v\rangle = \langle v, Av\rangle = \bar{\lambda}\langle v, v\rangle$. Par conséquent $\lambda = \bar{\lambda}$.

Le lemme de Schur réel 9.164 donne une matrice orthogonale qui trigonalise A. Les valeurs propres étant toutes réelles, la matrice QAQ^{-1} est même triangulaire (il n'y a pas de blocs dans la forme (9.365)). Prouvons que QAQ^{-1} est symétrique :

$$(QAQ^{-1})^t = (Q^{-1})^t A^t Q^t = QA^tQ^{-1} = QAQ^{-1} \tag{9.370}$$

où nous avons utilisé le fait que Q était orthogonale ($Q^{-1} = Q^t$) et que A était symétrique ($A^t = A$). Une matrice triangulaire supérieure symétrique est obligatoirement une matrice diagonale.

En ce qui concerne la base de vecteurs propres, soit $\{e_i\}_{i=1,\dots,n}$ la base canonique de $\mathbb{R}^n$ et Q une matrice orthogonale e telle que $A = Q^tDQ$ avec D diagonale. Nous posons $f_i = Q^te_i$ et en tenant compte du fait que $Q^t = Q^{-1}$ nous avons $Af_i = Q^tDQQ^te_i = Q^t\lambda_i e_i = \lambda_i f_i$. Donc les f_i sont des vecteurs propres de A. De plus ils sont orthonormés parce que

$$\langle f_i, f_j\rangle = \langle Q^te_i, Q^te_j\rangle = \langle e_i, Q^tQe_j\rangle = \langle e_i, e_j\rangle = \delta_{ij}. \tag{9.371}$$

□

Le théorème spectral pour les opérateurs auto-adjoints sera traité plus bas parce qu'il a besoin de choses sur les formes bilinéaires, théorème 9.261.

Remarque 9.166.
Une matrice symétrique est diagonalisable par une matrice orthogonale. Nous pouvons en réalité nous arranger pour diagonaliser par une matrice de $\mathrm{SO}(n)$. Plus généralement si A est une matrice diagonalisable par une matrice $P \in \mathrm{GL}^+(n,\mathbb{R})$ alors elle est diagonalisable par une matrice de $\mathrm{GL}^-(n,\mathbb{R})$ en changeant le signe de la première ligne de P. Et inversement.

En effet, si nous avons $P^tDP = A$, alors en notant $*$ les quantités qui ne dépendent pas de a, b ou c,

$$\begin{aligned}\begin{pmatrix} a & * & * \\ b & * & * \\ c & * & * \end{pmatrix}\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}\begin{pmatrix} a & b & c \\ * & * & * \\ * & * & * \end{pmatrix} &= \begin{pmatrix} a & * & * \\ b & * & * \\ c & * & * \end{pmatrix}\begin{pmatrix} \lambda_1 a & \lambda_1 b & \lambda_1 c \\ * & * & * \\ * & * & * \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 a^2 + * & \lambda_1 ab + * & \lambda_1 ac + * \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{pmatrix}.\end{aligned} \tag{9.372}$$

Nous voyons donc que si nous changeons les signes de a, b et c en même temps, le résultat ne change pas.

Définition 9.167 (Matrice définie positive, opérateur définit positif)**.**
Un opérateur sur un espace vectoriel sur $\mathbb{C}$ *ou* $\mathbb{R}$ *est* ***définit positif*** *si toutes ses valeurs propres sont réelles et strictement positives. Il est* ***semi-définie positive*** *si ses valeurs propres sont réelles positives ou nulles.*

Afin d'éviter l'une ou l'autre confusion, nous disons souvent *strictement* définie positive pour positive.

9.168.
Nous nommons $S^+(n,\mathbb{R})$ l'ensemble des matrices réelles symétriques $n \times n$ et $S^{++}(n,\mathbb{R})$ le sous-ensemble de $S^+(n,\mathbb{R})$ des matrices strictement définies positives.

Remarque 9.169.
Nous ne définissons pas la notion de matrice définie positive pour une matrice non symétrique.

Lorsqu'un énoncé parle d'une matrice symétrique, le premier réflexe est de la diagonaliser : considérer une matrice orthogonale T telle que $T^tMT = D$ avec D diagonale. Et les valeurs propres sur la diagonale : $D_{kl} = \delta_{kl}\lambda_k$. Les matrices symétriques définies positives ont cependant des propriétés même en dehors de leur base de diagonalisation.

Lemme 9.170.
Soit une matrice symétrique M.

(1) Elle est strictement définie positive si et seulement si $\langle x, Mx\rangle > 0$ pour tout x non nul dans $\mathbb{R}^n$.

(2) Elle est semi définie positive si et seulement si $\langle x, Mx\rangle \geqslant 0$ pour tout x non nul dans $\mathbb{R}^n$.

(3) Si elle est seulement définie positive, alors $\langle x, Mx\rangle \geqslant \lambda\|x\|^2$ dès que $\lambda \geqslant 0$ minore toutes les valeurs propres.

Démonstration. Démonstration en trois parties.

(1) Soit $\{e_i\}_{i=1,\dots,n}$ une base orthonormée de vecteurs propres de M dont l'existence est assurée par le théorème spectral 9.165. Nous nommons x_i les coordonnées de x dans cette base. Alors,

$$\langle x, Mx\rangle = \sum_{i,j} x_i\langle e_i, x_jMe_j\rangle = \sum_{i,j} x_ix_j\langle e_i, \lambda_je_j\rangle = \sum_{ij} x_ix_j\lambda_j\delta_{ij} = \sum_i \lambda_ix_i^2 \tag{9.373}$$

où les λ_i sont les valeurs propres de M. Cela est strictement positif pour tout x si et seulement si tous les λ_i sont strictement positifs.

(2) Nous avons encore

$$\langle x, Mx\rangle = \sum_i \lambda_ix_i^2. \tag{9.374}$$

Cela est plus grand ou égal à zéro si et seulement si tous les λ_i sont plus grands ou égaux à zéro.

(3) Soit une matrice orthogonale T diagonalisant M, c'est à dire telle que $T^tMT = D$ avec D diagonale. Nous allons vérifier que

$$\langle Tx, Mtx\rangle \geqslant \lambda\|Tx\|^2 \tag{9.375}$$

pour tout x. Vu que T est une bijection [39], cela impliquera le résultat pour tout x. Si nous considérons la base de diagonalisation $\{e_k\}$ pour les valeurs propres λ_k, nous avons le calcul

$$\begin{aligned}
\langle Tx, MTx\rangle &= \langle x, T^tMTx\rangle && \text{(9.376a)}\\
&= \langle x, Dx\rangle && \text{(9.376b)}\\
&= \sum_k \langle x, x_kDe_k\rangle && \text{(9.376c)}\\
&= \sum_k \lambda_kx_k\underbrace{\langle x, e_k\rangle}_{=x_k} && \text{(9.376d)}\\
&\geqslant \sum_k \lambda|x_k|^2 && \text{(9.376e)}\\
&= \lambda\|x\|^2 && \text{(9.376f)}\\
&= \lambda\|Tx\|^2. && \text{(9.376g)}
\end{aligned}$$

Au dernier passage nous avons utilisé le fait que T est une isométrie (proposition 9.50).

39. Une matrice orthogonale a un déterminant ± 1.

□

Les personnes qui aiment les vecteurs lignes et colonnes écriront des inégalités comme

$$x^tMx \geqslant x^tx. \tag{9.377}$$

Tout à l'autre bout du spectre des personnes névrosées des notations, on trouvera des inégalités comme

$$M(x \otimes x) \geqslant x \cdot x. \tag{9.378}$$

Le penchant personnel de l'auteur de ces lignes est la notation avec le produit tensoriel. Si vous aimez ça, vous pouvez lire 9.19.

La notation adoptée ici avec le produit scalaire $\langle x, Mx\rangle$ est entre les deux. Elle a l'avantage de n'être pas technologique comme le produit tensoriel (si vous y mettez les pieds, vous devez savoir ce que vous faites), tout en évitant de se casser la tête à savoir qui est un vecteur ligne ou un vecteur colonne.

Corollaire 9.171.
Une matrice symétrique strictement définie positive est inversible.

Démonstration. Si $Ax = 0$ alors $\langle Ax, x\rangle = 0$. Mais dans le cas d'une matrice strictement définie positive, cela implique $x = 0$ par le lemme 9.170. □

Lemme 9.172.
Pour une base quelconque, les éléments diagonaux d'une matrice symétrique semi-définie positive sont positifs. Si la matrice est strictement définie positive, alors les éléments diagonaux sont strictement positifs.

Démonstration. Il s'agit d'une application du lemme 9.170. Si A est définie positive et que $\{e_i\}$ est une base, alors

$$A_{ii} = \langle Ae_i, e_i\rangle \geqslant \lambda\|e_i\|^2 = \lambda \geqslant 0. \tag{9.379}$$

Si A est strictement définie positive, alors λ peut être choisit strictement positif. □

9.9.9 Pseudo-réduction simultanée

Corollaire 9.173 (Pseudo-réduction simultanée[108])**.**
Soient $A, B \in \mathrm{S}(n, \mathbb{R})$ avec A définie positive[40]*. Alors il existe $Q \in \mathrm{GL}(n, \mathbb{R})$ tell que Q^tBQ soit diagonale et $Q^tAQ = \mathbb{1}$.*

Démonstration. Nous allons noter $x \cdot y$ le produit scalaire usuel de $\mathbb{R}^n$ et $\{e_i\}_{i=1,\dots,n}$ sa base canonique.

Vu que A est définie positive, nous avons que l'expression[41] $\langle x, y\rangle = x \cdot Ay$ est un produit scalaire sur $\mathbb{R}^n$. Autrement dit, E muni de cette forme bilinéaire symétrique est un espace euclidien, ce qui fait dire à la proposition 9.40 qu'il existe une base de $\mathbb{R}^n$ orthonormée $\{f_i\}_{i=1,\dots,n}$ pour ce produit scalaire, c'est à dire qu'il existe une matrice $P \in \mathrm{GL}(n, \mathbb{R})$ telle que $P^tAP = \mathbb{1}$. Ici, P est la matrice de changement de base de la base canonique à notre base orthonormée, c'est à dire la matrice qui fait $Pe_i = f_i$ pour tout i. Voyons cela avec un peu de détails.

Pour savoir ce que valent les éléments de la matrice P^tAP, nous nous souvenons que P^tAPe_j est un vecteur dont les coordonnées sont les éléments de la j^{e}colonne de P^tAP. Nous avons donc $(P^tAP)_{ij} = e_i \cdot P^tAPe_i$. Calculons :

$$(P^tAP)_{ij} = e_i \cdot P^tAPe_i = Pe_i \cdot APe_j = f_i \cdot Af_j = \langle f_i, f_j\rangle = \delta_{ij} \tag{9.380}$$

où nous avons utilisé le fait que A était auto-adjointe pour la passer de l'autre côté du produit scalaire (usuel). Au final nous avons effectivement $P^tAP = \mathbb{1}$.

40. Définition 9.167.
41. On peut aussi l'écrire de façon plus matricielle sous la forme $\langle x, y\rangle = x^tAy$.

La matrice P^tBP est une matrice symétrique, donc le théorème spectral 9.165 nous donne une matrice $R \in \mathrm{O}(n,\mathbb{R})$ telle que R^tP^tBPR soit diagonale. En posant maintenant $Q = PR$ nous avons la matrice cherchée. □

Note : nous avons prouvé la pseudo-réduction simultanée comme corollaire du théorème de diagonalisation des matrices symétriques 9.165. Il aurait aussi pu être vu comme un corollaire du théorème spectral 9.261 sur les opérateurs auto-adjoints via son corollaire 9.262.

9.10 Sommes de familles infinies

9.10.1 Convergence commutative

Définition 9.174.
Soit x_k une suite dans un espace vectoriel normé E. Nous disons que la suite ***converge commutativement*** *vers $x \in E$ si $\lim_{n\to\infty} \|x_n - x\| = 0$ et si pour toute bijection $\tau\colon \mathbb{N} \to \mathbb{N}$ nous avons aussi*

$$\lim_{n\to\infty} \|x_{\tau(k)} - x\| = 0. \tag{9.381}$$

La notion de convergence commutative est surtout intéressante pour les séries. La somme

$$\sum_{k=0}^{\infty} x_k \tag{9.382}$$

converge commutativement vers x si $\lim_{N\to\infty} \|x - \sum_{k=0}^{N} x_k\| = 0$ et si pour toute bijection $\tau\colon \mathbb{N} \to \mathbb{N}$ nous avons

$$\lim_{N\to\infty} \|x - \sum_{k=0}^{N} x_{\tau(k)}\| = 0. \tag{9.383}$$

Nous démontrons maintenant qu'une série converge commutativement si et seulement si elle converge absolument.

Pour le sens inverse, nous avons la proposition suivante.

Proposition 9.175.
Soit $\sum_{k=0}^{\infty} a_k$ une série réelle qui converge mais qui ne converge pas absolument. Alors pour tout $b \in \mathbb{R}$, il existe une bijection $\tau\colon \mathbb{N} \to \mathbb{N}$ telle que $\sum_{i=0}^{\infty} a_{\tau(i)} = b$.

Pour une preuve, voir chez Gilles Dubois.

Proposition 9.176.
Soit $(a_i)_{i\in\mathbb{N}}$ une suite absolument convergente dans $\mathbb{C}$. Alors elle converge commutativement.

Démonstration. Soit $\epsilon > 0$. Nous posons $\sum_{i=0}^{\infty} a_i = a$ et nous considérons N tel que

$$|\sum_{i=0}^{N} a_i - a| < \epsilon. \tag{9.384}$$

Étant donné que la série des $|a_i|$ converge, il existe N_1 tel que pour tout $p, q > N_1$ nous ayons $\sum_{i=p}^{q} |a_i| < \epsilon$. Nous considérons maintenant une bijection $\tau\colon \mathbb{N} \to \mathbb{N}$. Prouvons que la série $\sum_{i=0}^{\infty} |a_{\tau(i)}|$ converge. Nous choisissons M de telle sorte que pour tout $n > M$, $\tau(n) > N_1$; alors si $p, q > M$ nous avons

$$\sum_{i=p}^{q} |a_{\tau(i)}| < \epsilon. \tag{9.385}$$

Par conséquent la somme de la suite $(a_{\tau(i)})$ converge. Nous devons montrer à présent qu'elle converge vers la même limite que la somme «usuelle» $\lim_{N\to\infty} \sum_{i=0}^{N} a_i$.

Soit $n > \max\{M, N\}$. Alors

$$\sum_{k=0}^{n} a_{\tau(k)} - \sum_{k=0}^{n} a_k = \sum_{k=0}^{M} a_{\tau(k)} - \sum_{k=0}^{N} a_k + \underbrace{\sum_{M+1}^{n} a_{\tau(k)}}_{<\epsilon} - \underbrace{\sum_{k=N+1}^{n} a_k}_{<\epsilon}. \tag{9.386}$$

Par construction les deux derniers termes sont plus petits que ϵ parce que M et N sont les constantes de Cauchy pour les séries $\sum a_{\tau(i)}$ et $\sum a_i$. Afin de traiter les deux premiers termes, quitte à redéfinir M, nous supposons que $\{1, \ldots, N\} \subset \tau\{1, \ldots, M\}$; par conséquent tous les a_i avec $i < N$ sont atteints par les $a_{\tau(i)}$ avec $i < M$. Dans ce cas, les termes qui restent dans la différence

$$\sum_{k=0} a_{\tau(k)} - \sum_{k=0}^{N} a_k \tag{9.387}$$

sont des a_k avec $k > N$. Cette différence est donc en valeur absolue plus petite que ϵ, et nous avons en fin de compte que

$$\left| \sum_{k=0}^{n} a_{\tau(k)} - \sum_{k=0}^{n} a_k \right| < \epsilon. \tag{9.388}$$

□

Proposition 9.177.
Soit $\sum_{i=0}^{\infty} a_i$ une série qui converge mais qui ne converge pas absolument. Pour tout $b \in \mathbb{R}$, il existe une bijection $\tau \colon \mathbb{N} \to \mathbb{N}$ telle que $\sum_{i=}^{\infty} a_{\tau(i)} = b$.

Les propositions 9.176 et 9.177 disent entre autres qu'une série dans $\mathbb{C}$ est commutativement sommable si et seulement si elle est absolument sommable.

Soit $(a_i)_{i\in I}$ une famille de nombres complexes indexée par un ensemble I quelconque. Nous allons nous intéresser à la somme $\sum_{i\in I} a_i$.

Soit $\{a_i\}_{i\in I}$ des nombres positifs. Nous définissons la somme

$$\sum_{i\in I} a_i = \sup_{J \text{ fini}} \sum_{j\in J} a_j. \tag{9.389}$$

Notons que cela est une définition qui ne fonctionne bien que pour les sommes de nombres positifs. Si $a_i = (-1)^i$, alors selon la définition nous aurions $\sum_i (-1)^i = \infty$. Nous ne voulons évidemment pas un tel résultat.

Dans le cas de familles de nombres réels positifs, nous avons une première définition de la somme.

Définition 9.178.
Soit $(a_i)_{i\in I}$ une famille de nombres réels positifs indexés par un ensemble quelconque I. Nous définissons

$$\sum_{i\in I} a_i = \sup_{J \textit{ fini dans } I} \sum_{j\in J} a_j. \tag{9.390}$$

Définition 9.179.
*Si $\{v_i\}_{i\in I}$ est une famille de vecteurs dans un espace vectoriel normé indexée par un ensemble quelconque I. Nous disons que cette famille est **sommable** de somme v si pour tout $\epsilon > 0$, il existe un J_0 fini dans I tel que pour tout ensemble fini K tel que $J_0 \subset K$ nous avons*

$$\| \sum_{j\in K} v_j - v \| < \epsilon. \tag{9.391}$$

Notons que cette définition implique la convergence commutative.

Exemple 9.180
La suite $a_i = (-1)^i$ n'est pas sommable parce que quel que soit J_0 fini dans $\mathbb{N}$, nous pouvons trouver

J fini contenant J_0 tel que $\sum_{j\in J}(-1)^j > 10$. Pour cela il suffit d'ajouter à J_0 suffisamment de termes pairs. De la même façon en ajoutant des termes impairs, on peut obtenir $\sum_{j\in J'}(-1)^i < -10$. △

Exemple 9.181
De temps en temps, la somme peut sortir d'un espace. Si nous considérons l'espace des polynômes $[0,1] \to \mathbb{R}$ muni de la norme uniforme, la somme de l'ensemble

$$\{1, -1, \pm\frac{x^n}{n!}\}_{n\in\mathbb{N}} \tag{9.392}$$

est zéro.

Par contre la somme de l'ensemble $\{1, \frac{x^n}{n!}\}_{n\in\mathbb{N}}$ est l'exponentielle qui n'est pas un polynôme. △

Exemple 9.182
Au sens de la définition 9.179 la famille

$$\frac{(-1)^n}{n} \tag{9.393}$$

n'est pas sommable. En effet la somme des termes pairs est ∞ alors que la somme des termes impairs est $-\infty$. Quel que soit $J_0 \in \mathbb{N}$, nous pouvons concocter, en ajoutant des termes pairs, un J avec $J_0 \subset J$ tel que $\sum_{j\in J}(-1)^j/j$ soit arbitrairement grand. En ajoutant des termes négatifs, nous pouvons également rendre $\sum_{j\in J}(-1)^j/j$ arbitrairement petit. △

Proposition 9.183.
Si (a_{ij}) est une famille de nombres positifs indexés par $\mathbb{N}\times\mathbb{N}$ alors

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} = \sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty} a_{ij}\Big) \tag{9.394}$$

où la somme de gauche est celle de la définition 9.178.

Démonstration. Nous considérons $J_{m,n} = \{0,\dots,m\}\times\{0,\dots,n\}$ et nous avons pour tout m et n :

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} \geqslant \sum_{(i,j)\in J_{m,n}} a_{ij} = \sum_{i=1}^{m}\Big(\sum_{j=1}^{n} a_{ij}\Big). \tag{9.395}$$

Si nous fixons m et que nous prenons la limite $n\to\infty$ (qui commute avec la somme finie sur i) nous trouvons

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} \geqslant= \sum_{i=1}^{m}\Big(\sum_{j=1}^{\infty} a_{ij}\Big). \tag{9.396}$$

Cela étant valable pour tout m, c'est encore valable à la limite $m\to\infty$ et donc

$$\sum_{(i,j)\in\mathbb{N}^2} a_{ij} \geqslant \sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty} a_{ij}\Big). \tag{9.397}$$

Pour l'inégalité inverse, il faut remarquer que si J est fini dans $\mathbb{N}^2$, il est forcément contenu dans $J_{m,n}$ pour m et n assez grand. Alors

$$\sum_{(i,j)\in J} a_{ij} \leqslant \sum_{(i,j)\in J_{m,n}} a_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} \leqslant \sum_{i=1}^{\infty}\Big(\sum_{j=1}^{\infty} a_{ij}\Big). \tag{9.398}$$

Cette inégalité étant valable pour tout ensemble fini $J\subset\mathbb{N}^2$, elle reste valable pour le supremum. □

La définition générale de la somme 9.179 est compatible avec la définition usuelle dans les cas où cette dernière s'applique.

Proposition 9.184 (commutative sommabilité)**.**
Soit I un ensemble dénombrable et une bijection $\tau\colon \mathbb{N} \to I$. Soit $(a_i)_{i\in I}$ une famille dans un espace vectoriel normé. Alors

$$\sum_{k=0}^{\infty} a_{\tau(k)} = \sum_{i\in I} a_i \tag{9.399}$$

dès que le membre de droite existe. Le membre de gauche est définit par la limite usuelle.

Démonstration. Nous posons $a = \sum_{i\in I} a_i$. Soit $\epsilon > 0$ et J_0 comme dans la définition. Nous choisissons

$$N > \max_{j\in J_0}\{\tau^{-1}(j)\}. \tag{9.400}$$

En tant que sommes sur des ensembles finis, nous avons l'égalité

$$\sum_{k=0}^{N} a_{\tau(k)} = \sum_{j\in J_0} a_j \tag{9.401}$$

où J est un sous-ensemble de I contenant J_0. Soit J fini dans I tel que $J_0 \subset J$. Nous avons alors

$$\| \sum_{k=0}^{N} a_{\tau(k)} - a\| = \| \sum_{j\in J} a_j - a\| < \epsilon. \tag{9.402}$$

Nous avons prouvé que pour tout ϵ, il existe N tel que $n > N$ implique $\| \sum_{k=0}^{n} a_{\tau(k)} - a\| < \epsilon$. □

Corollaire 9.185.
Nous pouvons permuter une somme dénombrable et une fonction linéaire continue. C'est à dire que si f est une fonction linéaire continue sur l'espace vectoriel normé E et $(a_i)_{i\in I}$ une famille sommable dans E alors

$$f\left(\sum_{i\in I} a_i\right) = \sum_{i\in I} f(a_i). \tag{9.403}$$

Démonstration. En utilisant une bijection τ entre I et $\mathbb{N}$ avec la proposition 9.184 ainsi que le résultat connu à propos des sommes sur $\mathbb{N}$, nous avons

$$f\left(\sum_{i\in I} a_i\right) = f\left(\sum_{k=0}^{\infty} a_{\tau(k)}\right) \tag{9.404a}$$

$$= \sum_{k=0}^{\infty} f(a_{\tau(k)}) \tag{9.404b}$$

$$= \sum_{i\in I} f(a_i). \tag{9.404c}$$

Notons que le passage à (9.404b) n'est pas du tout une trivialité à deux francs cinquante. Il s'agit d'écrire la somme comme la limite des sommes partielles, et de permuter f avec la limite en invoquant la continuité, puis de permuter f avec la somme partielle en invoquant sa linéarité.

Ah, tiens et tant qu'on y est à dire qu'il y a des chose évidentes qui ne le sont pas, oui, il existe des applications linéaires non continues, voir le thème 19. □

La proposition suivante nous enseigne que les sommes infinies peuvent être manipulée de façon usuelle.

Proposition 9.186.
Soit I un ensemble dénombrable. Soient $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$, deux familles de réels positifs telles que $a_i < b_i$ et telles que (b_i) est sommable. Alors (a_i) est sommable.

Si $(a_i)_{i\in I}$ est une famille de complexes telle que $(|a_i|)$ est sommable, alors (a_i) est sommable.

Proposition 9.187 ([1])**.**
Soit un espace vectoriel normé E et une famille sommable[42] $\{v_i\}_{i\in I}$ *d'éléments de E. Soit $f\colon E\to\mathbb{C}$ une application sur laquelle nous supposons*

(1) f est linéaire et continue ;

(2) la partie $\{f(v_i)_{i\in I}\}$ est sommable.

Alors nous pouvons permuter la somme et f :

$$f\big(\sum_{i\in I} v_i\big)=\sum_{i\in I} f(v_i). \tag{9.405}$$

Démonstration. Soit $\epsilon>0$; vu que les familles $\{v_i\}_{i\in I}$ et $\{f(v_i)\}_{i\in I}$ sont sommables, nous pouvons considérer les parties finies J_1 et J_2 de I telles que

$$\Big\|\sum_{j\in J_1} v_j-\sum_{i\in I} v_i\Big\|\leqslant\epsilon \tag{9.406}$$

et

$$\Big\|\sum_{j\in J_2} f(v_j)-\sum_{i\in I} f(v_i)\Big\|\leqslant\epsilon \tag{9.407}$$

Ensuite nous posons $J=J_1\cup J_2$. Avec cela nous calculons un peu avec les majorations usuelles :

$$\|f(\sum_{i\in I} v_i)-\sum_{i\in I} f(v_i)\|\leqslant\|f(\sum_{i\in I} v_i)-f(\sum_{j\in J} v_j)\|+\|f(\sum_{j\in J} v_j)-\sum_{i} \in I f(v_i)\|. \tag{9.408}$$

Le second terme est majoré par ϵ, tandis que le premier, en utilisant la linéarité de f possède la majoration

$$\|f(\sum_{i\in I} v_i)-f(\sum_{j\in J} v_j)\|=\|f(\sum_{i\in I} v_i-\sum_{j\in J} v_j)\|\leqslant\|f\|\|\sum_{i\in I} v_i-\sum_{j\in J} v_j\|\leqslant\epsilon\|f\|. \tag{9.409}$$

Donc pour tout $\epsilon>0$ nous avons

$$\|f(\sum_{i\in I} v_i)-\sum_{i\in I} f(v_i)\|\leqslant\epsilon(1+\|f\|). \tag{9.410}$$

D'où l'égalité (9.405). □

9.11 Fonctions

Soient $(V,\|.\|_V)$ et $(W,\|.\|_W)$ deux espaces vectoriels normés, et une fonction f de V dans W. Il est maintenant facile de définir les notions de limites et de continuité pour de telles fonctions en copiant les définitions données pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ en changeant simplement les valeurs absolues par les normes sur V et W.

La caractérisation suivante est un recopiage de la définition 6.69 lorsque la topologie est donnée par des boules.

Proposition 9.188.
Soit $f\colon V\to W$ une fonction de domaine $\mathrm{Dom}(f)\subset V$ et soit a un point d'accumulation de $\mathrm{Dom}(f)$. La fonction f admet une limite en $a\in V$ si et seulement s'il existe un élément $\ell\in W$ tel que pour tout $\varepsilon>0$, il existe un $\delta>0$ tel que pour tout $x\in\mathrm{Dom}(f)$,

$$0<\|x-a\|_V<\delta\ \Rightarrow\ \|f(x)-\ell\|_W<\varepsilon. \tag{9.411}$$

*Dans ce cas, nous écrivons $\lim_{x\to a} f(x)=\ell$ et nous disons que ℓ est la **limite** de f lorsque x tend vers a.*

42. Définition 9.179.

Remarque 9.189.
Le fait que nous limitions la formule (9.411) aux x dans le domaine de f n'est pas anodin. Considérons la fonction $f(x) = \sqrt{x^2 - 4}$, de domaine $|x| \geqslant 2$. Nous avons

$$\lim_{x \to 2} \sqrt{x^2 - 4} = 0. \tag{9.412}$$

Nous ne pouvons pas dire que cette limite n'existe pas en justifiant que la limite à gauche n'existe pas. Les points $x < 2$ sont hors du domaine de f et ne comptent dons pas dans l'appréciation de l'existence de la limite.

Vous verrez plus tard que ceci provient de la topologie induite de $\mathbb{R}$ sur l'ensemble $[2, \infty[$.

9.12 Sous espaces caractéristiques

Lorsqu'un opérateur n'est pas diagonalisable, les valeurs propres jouent quand même un rôle important.

Définition 9.190.
Soit E un $\mathbb{K}$-espace vectoriel $f \in \mathrm{End}(E)$. Pour $\lambda \in \mathbb{K}$ nous définissons

$$F_\lambda(f) = \{v \in E \text{ tel que } (f - \lambda \mathbb{1})^n v = 0, n \in \mathbb{N}\} \tag{9.413}$$

*et nous appelons ça un **sous-espace caractéristique** de f.*

L'espace $F_\lambda(f)$ est l'ensemble de nilpotence de l'opérateur $f - \lambda\mathbb{1}$ et

Lemme 9.191.
L'ensemble $F_\lambda(f)$ est non vide si et seulement si λ est une valeur propre de f. L'espace $F_\lambda(f)$ est invariant sous f.

Démonstration. Si $F_\lambda(f)$ est non vide, nous considérons $v \in F_\lambda(f)$ et n le plus petit entier non nul tel que $(f - \lambda)^n v = 0$. Alors $(f - \lambda)^{n-1} v$ est un vecteur propre de f pour la valeur propre λ. Inversement si v est une valeur propre de f pour la valeur propre λ, alors $v \in F_\lambda(f)$.

En ce qui concerne l'invariance, remarquons que f commute avec $f - \lambda\mathbb{1}$. Si $x \in F_\lambda(f)$ il existe n tel que $(f - \lambda\mathbb{1})^n x = 0$. Nous avons aussi

$$(f - \lambda\mathbb{1})^n f(x) = f\big((f - \lambda\mathbb{1})^n x\big) = 0, \tag{9.414}$$

par conséquent $f(x) \in F_\lambda(f)$. □

Remarque 9.192.
Toute matrice sur $\mathbb{C}$ n'est pas diagonalisable : nous en avons déjà donné une exemple simple en 9.239. Nous en voyons maintenant un moins simple. Considérons en effet l'endomorphisme f donné par la matrice

$$\begin{pmatrix} a & \alpha & \beta \\ 0 & a & \gamma \\ 0 & 0 & b \end{pmatrix} \tag{9.415}$$

où $a \neq b$, $\alpha \neq 0$, β et γ sont des nombres complexes quelconques. Son polynôme caractéristique est

$$\chi_f(\lambda) = (a - \lambda)^2 (b - \lambda), \tag{9.416}$$

et les valeurs propres sont donc a et b. Nous trouvons les vecteurs propres pour la valeur a en résolvant

$$\begin{pmatrix} a & \alpha & \beta \\ 0 & a & \gamma \\ 0 & 0 & b \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} ax \\ ay \\ az \end{pmatrix}. \tag{9.417}$$

L'espace propre $E_a(f)$ est réduit à une seule dimension générée par $(1,0,0)$. De la même façon l'espace propre correspondant à la valeur propre b est donné par

$$\begin{pmatrix} \frac{1}{b-a}\left(\beta+\frac{\alpha\gamma}{b-a}\right) \\ \frac{\gamma}{b-a} \\ 1 \end{pmatrix}. \tag{9.418}$$

Il n'y a donc pas trois vecteurs propres linéairement indépendants, et l'opérateur f n'est pas diagonalisable.

Par contre nous pouvons voir que

$$(f-\alpha\mathbb{1})^2\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}a&\alpha&\beta\\0&a&\gamma\\0&0&b\end{pmatrix}\begin{pmatrix}\alpha\\0\\0\end{pmatrix}-\begin{pmatrix}a\alpha\\0\\0\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}, \tag{9.419}$$

de telle sorte que le vecteur $(0,1,0)$ soit également dans l'espace caractéristique $F_a(f)$.

Dans cet exemple, la multiplicité algébrique de la racine a du polynôme caractéristique vaut 2 tandis que sa multiplicité géométrique vaut seulement 1.

9.12.1 Théorèmes de décomposition

Théorème 9.193 (Théorème spectral, décomposition primaire)**.**
Soit E espace vectoriel de dimension finie sur le corps algébriquement clos $\mathbb{K}$ et $f \in \mathrm{End}(E)$. Alors

$$E = F_{\lambda_1}(f) \oplus \ldots \oplus F_{\lambda_k}(f) \tag{9.420}$$

où la somme est sur les valeurs propres distinctes de f.

Les projecteurs sur les espaces caractéristique forment un système complet et orthogonal.

Démonstration. Soit P le polynôme caractéristique de f et une décomposition

$$P = (f-\lambda_1)^{\alpha_1}\ldots(f-\lambda_r)^{\alpha_r} \tag{9.421}$$

en facteurs irréductibles. La le théorème de noyaux (9.102) nous avons

$$E = \ker(f-\lambda_1)^{\alpha_1} \oplus \ldots \oplus \ker(f-\lambda_r)^{\alpha_r}. \tag{9.422}$$

Les projecteurs sont des polynômes en f et forment un système orthogonal. Il nous reste à prouver que $\ker(f-\lambda_i)^{\alpha_i} = F_{\lambda_i}(f)$. L'inclusion

$$\ker(f-\lambda_i)^{\alpha_i} \subset F_{\lambda_i}(f) \tag{9.423}$$

est évidente. Nous devons montrer l'inclusion inverse. Prouvons que la somme des $F_{\lambda_i}(f)$ est directe. Si $v \in F_{\lambda_i}(f) \cap F_{\lambda_j}(f)$, alors il existe $v_1 = (f-\lambda_i)^n v \neq 0$ avec $(f-\lambda_i)v_1 = 0$. Étant donné que $(f-\lambda_i)$ commute avec $(f-\lambda_j)$, ce v_1 est encore dans $F_{\lambda_j}(f)$ et par conséquent il existe $w = (f-\lambda_j)^m v_1$ non nul tel que

$$\begin{cases} (f-\lambda_i)w = 0 & \text{(9.424a)} \\ (f-\lambda_j)w = 0. & \text{(9.424b)} \end{cases}$$

Ce w serait donc un vecteur propre simultané pour les valeurs propres λ_i et λ_j, ce qui est impossible parce que les espaces propres sont linéairement indépendants. Les espaces F_{λ_i} sont donc en somme directe et

$$\sum_i \dim F_{\lambda_i}(f) \leqslant \dim E. \tag{9.425}$$

En tenant compte de l'inclusion (9.423) nous avons même

$$\dim E = \sum_i \dim\ker(f-\lambda_i)^{\alpha_i} \leqslant \sum_i F_{\lambda_i}(f) \leqslant \dim E. \tag{9.426}$$

Par conséquent nous avons $\dim\ker(f-\lambda_i)^{\alpha_i} = \dim F_{\lambda_i}(f)$ et l'égalité des deux espaces. □

Problèmes et choses à faire

Dans le cas où le corps n'est pas algébriquement clos, il paraît qu'il faut remplacer «diagonalisable» par «semi-simple».

Si l'espace vectoriel est sur un corps algébriquement clos, alors les endomorphismes semi-simples[43] sont les endomorphismes diagonaux.

Théorème 9.194 (Décomposition de Dunford)**.**
Soit E un espace vectoriel sur le corps algébriquement clos $\mathbb{K}$ et $u \in \mathrm{End}(E)$ un endomorphisme de E.

(1) L'endomorphisme u se décompose de façon unique sous la forme

$$u = s + n \tag{9.427}$$

où s est diagonalisable, n est nilpotent et $[s, n] = 0$.

(2) Les endomorphismes s et n sont des polynômes en u et commutent avec u.

(3) Les parties s et n sont données par

$$s = \sum_i \lambda_i p_i \tag{9.428a}$$

$$n = \sum_i (s - \lambda_i \mathbb{1}) p_i \tag{9.428b}$$

où les sommes sont sur les valeurs propres distinctes[44] de f et où $p_i \colon E \to F_{\lambda_i}(u)$ est la projection de E sur $F_{\lambda_i}(u)$.

Démonstration. Le théorème spectral 9.193 nous indique que

$$E = \bigoplus_i F_{\lambda_i}(f). \tag{9.429}$$

Nous considérons l'endomorphisme s de E qui consiste à dilater d'un facteur λ l'espace caractéristique $F_\lambda(f)$:

$$s = \sum_i \lambda_i p_i \tag{9.430}$$

où $p_i \colon E \to F_{\lambda_i}(u)$ est la projection de E sur $F_{\lambda_i}(u)$.

Nous allons prouver que $[s, f] = 0$ et $n = f - s$ est nilpotent. Cela impliquera que $[s, n] = 0$.

Si $x \in F_\lambda(f)$, alors nous avons $sf(x) = \lambda f(x)$ parce que $f(x) \in F_\lambda(f)$ tandis que $fs(x) = f(\lambda x) = \lambda f(x)$. Par conséquent f commute avec s.

Pour montrer que $f - s$ est nilpotent, nous en considérons la restriction

$$f - s \colon F_\lambda(f) \to F_\lambda(f). \tag{9.431}$$

Cet opérateur est égal à $f - \lambda\mathbb{1}$ et est par conséquent nilpotent.

Prouvons à présent l'unicité. Soit $u = s' + n'$ une autre décomposition qui satisfait aux conditions : s' est diagonalisable, n' est nilpotent et $[n', s'] = 0$. Commençons par prouver que s' et n' commutent avec u. En multipliant $u = s' + n'$ par s' nous avons

$$s'u = s'^2 + s'n' = s'^2 + n's' = (s' + n')s' = us', \tag{9.432}$$

par conséquent $[u, s'] = 0$. Nous faisons la même chose avec n' pour trouver $[u, n'] = 0$. Notons que pour obtenir ce résultat nous avons utilisé le fait que n' et s' commutent, mais pas leur propriétés de nilpotence et de diagonalisibilité.

Si $s' + n' = s + n$ est une autre décomposition, s' et n' commutent avec u, et par conséquent avec tous les polynômes en u. Ils commutent en particulier avec n et s. Les endomorphismes s

43. Définition 9.116.
44. C'est à dire sur les sous-espaces caractéristiques.

et s' sont alors deux endomorphismes diagonalisables qui commutent. Par la proposition 9.146, ils sont simultanément diagonalisables. Dans la base de simultanée diagonalisation, la matrice de l'opérateur $s' - s = n - n'$ est donc diagonale. Mais $n - n'$ est également nilpotent, en effet si A et B sont deux opérateurs nilpotents,

$$(A + B)^n = \sum_{k=0}^{n} \binom{k}{n} A^k B^{n-k}. \tag{9.433}$$

Si n est assez grand, au moins un parmi A^k ou B^{n-k} est nul.

Maintenant que $n - n'$ est diagonal et nilpotent, il est nul et $n = n'$. Nous avons alors immédiatement aussi $s = s'$.

□

9.12.2 Diverses conséquences

Théorème 9.195.
Soit une matrice $A \in \mathbb{M}(n, \mathbb{C})$. On a que la suite $(A^k x)$ tends vers zéro pour tout x si et seulement si $\rho(A) < 1$ où $\rho(A)$ est le rayon spectral de A

Démonstration. Dans le sens direct, il suffit de prendre comme x, un vecteur propre de A. Dans ce cas nous avons $A^k x = \lambda^k x$. Mais $\lambda^k x$ ne tend vers zéro que si $\lambda < 1$. Donc toute les valeurs propres de A doivent être plus petite que 1 et $\rho(A) < 1$.

Pour l'autre sens nous utilisons la décomposition de Dunford (théorème 9.194) : il existe une matrice inversible P telle que

$$A = P^{-1}(D + N)P \tag{9.434}$$

où D est diagonale, N est nilpotente et $[D, N] = 0$. Étant donné que $D + N$ est triangulaire, son polynôme caractéristique que

$$\chi_{D+N}(\lambda) = \prod_i D_{ii} - \lambda. \tag{9.435}$$

Par similitude, c'est le même polynôme caractéristique que celui de A et nous savons alors que la diagonale de D contient les valeurs propres de A.

Par ailleurs nous avons

$$A^k = P^{-1}(D + N)^k P \tag{9.436a}$$

$$= P^{-1} \sum_{j=0}^{k} \binom{j}{k} D^{j-k} N^j P \tag{9.436b}$$

$$= P^{-1} \sum_{j=0}^{n-1} \binom{j}{k} D^{j-k} N^j P \tag{9.436c}$$

où nous avons utilité le fait que D et N commutent ainsi que $N^{n-1} = 0$ parce que N est nilpotente. Nous utilisons la norme matricielle usuelle, pour laquelle $\|D\| = \rho(D) = \rho(A)$. Nous avons alors

$$\|(D + N)^k\| \leqslant \sum_{j=0}^{k} \binom{j}{k} \rho(D)^{k-j} \|N\|^j. \tag{9.437}$$

Du coup si $\rho(D) < 1$ alors $\|(D + N)^k\| \to 0$ (et c'est même un si et seulement si). □

Une application de la décomposition de Jordan est l'existence d'un logarithme pour les matrices. La proposition suivant va d'une certaine manière donner un logarithme pour les matrices inversibles complexes. Dans le cas des matrices réelles m telles que $\|m - \mathbb{1}\| < 1$, nous donnerons au lemme 15.92 une formule pour le logarithme sous forme d'une série ; ce logarithme sera réel.

Proposition 9.196.
Toute matrice inversible complexe est une exponentielle.

Démonstration. Soit $A \in \mathrm{GL}(n,\mathbb{C})$; nous allons donner une matrice $B \in \mathbb{M}(n,\mathbb{C})$ telle que $A = \exp(B)$. D'abord remarquons qu'il suffit de prouver le résultat pour une matrice par classe de similitude. En effet si $A = \exp(B)$ et si M est inversible alors

$$\exp(MBM^{-1}) = \sum_k \frac{1}{k!}(MBM^{-1})^k \tag{9.438a}$$

$$= \sum_k \frac{1}{k!} MB^k M^{-1} \tag{9.438b}$$

$$= M\exp(B)M^{-1}. \tag{9.438c}$$

Donc $MAM^{-1} = \exp(MBM^{-1})$. Nous pouvons donc nous contenter de trouver un logarithme pour les blocs de Jordan. Nous supposons donc que $A = (\mathbb{1} + N)$ avec $N^m = 0$. En nous inspirant de (15.172), nous posons [45]

$$D(t) = tN - \frac{t^2}{2}N^2 + \cdots + (-1)^m \frac{t^{m-1}}{m-1}N^{m-1} \tag{9.439}$$

et nous allons prouver que $e^{D(1)} = \mathbb{1} + N$. Notons que N étant nilpotente, cette somme ainsi que toutes celles qui viennent sont finies. Il n'y a donc pas de problèmes de convergences dans cette preuve (si ce n'est les passages des équations (9.438)).

Nous posons $S(t) = e^{D(t)}$ (la somme est finie), et nous avons

$$S'(t) = D'(t)e^{D(t)} \tag{9.440}$$

Afin d'obtenir une expression qui donne S' en termes de S, nous multiplions par $(\mathbb{1} + tN)$ en remarquant que $(\mathbb{1} + tN)D'(t) = N$ nous avons

$$(\mathbb{1} + tN)S'(t) = NS(t). \tag{9.441}$$

En dérivant à nouveau,

$$(\mathbb{1} + tN)S''(t) = 0. \tag{9.442}$$

La matrice $(\mathbb{1} + tN)$ est inversible parce que son noyau est réduit à $\{0\}$. En effet si $(\mathbb{1} + tN)x = 0$, alors $Nx = -\frac{1}{t}x$, ce qui est impossible parce que N est nilpotente. Ce que dit l'équation (9.442) est alors que $S''(t) = 0$. Si nous développons $S(t)$ en puissances de t nous nous arrêtons au terme d'ordre 1 et nous avons

$$S(t) = S(0) + tS'(0) = \mathbb{1} + tD'(0) = 1 + tN. \tag{9.443}$$

En $t = 1$ nous trouvons $S(1) = \mathbb{1} + N$. La matrice $D(1)$ donnée est donc bien un logarithme de $\mathbb{1} + N$. □

9.12.3 Diagonalisabilité d'exponentielle

Proposition 9.197 ([34])**.**
Si $A \in \mathbb{M}(n,\mathbb{R})$ a un polynôme caractéristique scindé, alors A est diagonalisable si et seulement si e^A est diagonalisable.

Démonstration. Si A est diagonalisable, alors il existe une matrice inversible M telle que $D = M^{-1}AM$ soit diagonale (c'est la définition 9.140). Dans ce cas nous avons aussi $(M^{-1}AM)^k = M^{-1}A^kM$ et donc $M^{-1}e^AM = e^{M^{-1}AM} = e^D$ qui est diagonale.

La partie difficile est donc le contraire.

45. Le logarithme d'un nombre n'est pas encore définit à ce moment, mais cela ne nous empêche pas de poser une définition ici pour une application des réels vers les matrices.

Qui est diagonalisable et comment ? Nous supposons que e^A est diagonalisable et nous écrivons la décomposition de Dunford (théorème 9.194) :

$$A = S + N \tag{9.444}$$

où S est diagonalisable, N est nilpotente, $[S, N] = 0$. Nous avons besoin de prouver que $N = 0$.

Les matrices A est S commutent ; en passant au développement nous en déduisons que A et e^S commutent, puis encore en passant au développement que e^A et e^S commutent. Vu que S est diagonalisable, e^S l'est et par hypothèse e^A est également diagonalisable. Donc e^A et e^{-S} sont simultanément diagonalisables par la proposition 9.146.

Étant donné que A et S commutent, nous avons $e^N = e^{A-S} = e^A e^{-S}$, et nous en déduisons que e^N est diagonalisable vu que les deux facteurs e^A et e^{-S} sont simultanément diagonalisables.

Unipotence Si r est le degré de nilpotence de N, nous avons

$$e^N - \mathbb{1} = N + \frac{N^2}{2} + \cdots + \frac{N^{r-1}}{(r-1)!}. \tag{9.445}$$

Donc

$$(e^N - \mathbb{1})^k = \left(N + \frac{N^2}{2} + \cdots + \frac{N^{r-1}}{(r-1)!} \right)^k \tag{9.446}$$

où le membre de droite est un polynôme en N dont le terme de plus bas degré est de degré k. Donc $(e^N - \mathbb{1})$ est nilpotente et e^N est unipotente.

Si M est la matrice qui diagonalise e^N, alors la matrice diagonale $M^{-1}e^N M$ est tout autant unipotente que e^N elle-même. En effet,

$$(M^{-1}e^N M - \mathbb{1})^r = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} M^{-1} (e^N)^k M \tag{9.447a}$$

$$= M^{-1} \left(\sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} (e^N)^k \right) M \tag{9.447b}$$

$$= M^{-1}(e^N - \mathbb{1})^r M \tag{9.447c}$$

$$= 0. \tag{9.447d}$$

La matrice $M^{-1}e^N M$ est donc une matrice diagonale et unipotente ; donc $M^{-1}e^N M = \mathbb{1}$, ce qui donne immédiatement que $e^N = \mathbb{1}$.

Polynômes annulateurs En reprenant le développement (9.445) sachant que $e^N = \mathbb{1}$, nous savons que

$$N + \frac{N^2}{2} + \cdots + \frac{N^{r-1}}{(r-1)!} = 0. \tag{9.448}$$

Dit en termes pompeux (mais non moins porteurs de sens), le polynôme

$$Q(X) = X + \frac{X^2}{2} + \cdots + \frac{X^{r-1}}{(r-1)!} \tag{9.449}$$

est un polynôme annulateur de N.

La proposition 9.110 stipule que le polynôme minimal d'un endomorphisme divise tous les polynômes annulateurs. Dans notre cas, X^r est un polynôme annulateur et donc le polynôme minimal de N est de la forme X^k. Donc il est X^r lui-même.

Nous avons donc $X^r \mid Q$. Mais Q est un polynôme contenant le monôme X donc X^r ne peut diviser Q que si $r = 1$. Nous en concluons que X est un polynôme annulateur de N. C'est à dire que $N = 0$.

Conclusion Vu que Dunford[46] dit que $A = S + N$ et que nous venons de prouver que $N = 0$, nous concluons que $A = S$ avec S diagonalisable.

□

46. Théorème 9.194.

9.12.4 Valeurs singulières

Définition 9.198.
*Soit M une matrice $m \times n$ sur $\mathbb{K}$ ($\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$). Un nombre réel σ est une **valeur singulière** de M s'il existent des vecteurs unitaires $u \in \mathbb{K}^m$, $v \in \mathbb{K}^n$ tels que*

$$Mv = \sigma u \tag{9.450a}$$
$$M^*u = \sigma v. \tag{9.450b}$$

Théorème 9.199 (Décomposition en valeurs singulières)**.**
Soit $M \in \mathbb{M}(m \times n, \mathbb{K})$ où $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Alors M se décompose en

$$M = ADB \tag{9.451}$$

où il existe deux matrices unitaires $A \in \mathrm{U}(m \times m)$, $B \in \mathrm{U}(n \times n)$ et une matrice (pseudo)diagonale $D \in \mathbb{M}(m \times n)$ tels que

(1) $A \in \mathrm{U}(m \times m)$, $B \in \mathrm{U}(n \times n)$ sont deux matrices unitaires ;,

(2) D est (pseudo)diagonale,

(3) les éléments diagonaux de D sont les valeurs singulières de M,

(4) le nombre d'éléments non nuls sur la diagonale de D est le rang de M.

Corollaire 9.200.
Soit $M \in \mathbb{M}(n, \mathbb{C})$. Il existe un isomorphisme $f : \mathbb{C}^n \to \mathbb{C}^n$ tel que fM soit autoadjoint.

Démonstration. Si $M = ADB$ est la décomposition de M en valeurs singulières, alors nous pouvons prendre $f = \overline{B}^t A^{-1}$ qui est une matrice inversible. Pour la vérification que ce f répond bien à la question, ne pas oublier que D est réelle, même si M ne l'est pas. □

9.13 Extension du corps de base

Nous avons discuté dans la section 5.4 de ce qui arrive au corps lorsqu'on l'étend. Dans cette sections nous allons étudier ce qui arrive aux applications linéaires entre deux $\mathbb{K}$-espaces vectoriels lorsque nous étendons le corps $\mathbb{K}$ en un corps $\mathbb{L}$.

Soit donc un corps $\mathbb{K}$ et deux $\mathbb{K}$-espaces vectoriels E et F, et entrons dans le vif du sujet[47]. Soit $\mathbb{K}$ un corps (commutatif) et une extension $\mathbb{L}$ de $\mathbb{K}$. Soient E et F, des $\mathbb{K}$-espaces vectoriels de dimension finie.

9.13.1 Extension des applications linéaires

Définition 9.201 ([109])**.**
*L'espace vectoriel obtenu par **extension du corps de base** de E est l'espace vectoriel*

$$E_{\mathbb{L}} = \mathbb{L} \otimes_{\mathbb{K}} E. \tag{9.452}$$

Ce dernier est le quotient $\mathbb{L} \otimes_{\mathbb{K}} E = (\mathbb{L} \times E)/\sim$ par la relation d'équivalence

$$(\lambda, v) \sim \big(a\lambda, \frac{1}{a}v\big) \tag{9.453}$$

pour tout $a \in \mathbb{K}$. Nous noterons $[\lambda, v]$ ou $\lambda \otimes v$ ou encore $\lambda \otimes_{\mathbb{K}} v$ la classe de (λ, v).

Un élément de $E_{\mathbb{L}}$ est de la forme $\sum_k [\lambda_k, v_k]$ avec $\lambda_k \in \mathbb{L}$ et $v_k \in E$. Si $f : E \to F$ est une applications linéaire nous définissons

$$\begin{aligned} f_{\mathbb{L}} : E_{\mathbb{L}} &\to F_{\mathbb{L}} \\ [\lambda, v] &\mapsto [\lambda, f(v)]. \end{aligned} \tag{9.454}$$

47. Le sujet étant le corps étendu.

Remarque 9.202.
Si deux vecteurs de $E_{\mathbb{L}}$ sont linéairement indépendants pour $\mathbb{K}$, ils ne le sont pas spécialement pour $\mathbb{L}$. Par exemple si $\mathbb{C}$ est vu comme $\mathbb{R}$-espace vectoriel, alors $\{1, i\}$ est une partie libre. Mais dans $\mathbb{C}$ vu comme $\mathbb{C}$-espace vectoriel, la partie $\{1, i\}$ n'est pas libre.

Nous définissons aussi l'injection canonique

$$\begin{aligned} \iota\colon E &\to E_{\mathbb{L}} \\ v &\mapsto [1, v]. \end{aligned} \tag{9.455}$$

Proposition 9.203 ([83])**.**
Injectivité et surjectivité respectées.

(1) L'application $f_{\mathbb{L}}$ est injective si et seulement si f est injective.

(2) L'application $f_{\mathbb{L}}$ est surjective si et seulement si f est surjective.

Démonstration. Supposons pour commencer que $f_{\mathbb{L}}$ est injective. Le diagramme

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\tau} \\ E_{\mathbb{L}} & \xrightarrow[f_{\mathbb{L}}]{} & F_{\mathbb{L}} \end{array} \tag{9.456}$$

est un diagramme commutatif. En effet

$$(\tau \circ f)(v) = [1, f(v)] \tag{9.457}$$

tandis que

$$(f_{\mathbb{L}\circ\tau})(v) = f_{\mathbb{L}}[1, v] = [1, f(v)]. \tag{9.458}$$

Donc si $f(v) = 0$ avec $v \neq 0$ nous aurions $(\tau \circ f)(v) = 0$ et donc aussi $(f_{\mathbb{L}} \circ \tau)(v) = 0$, alors que $\tau(v) \neq 0$ dans $E_{\mathbb{L}}$.

Réciproquement, nous supposons que f est injective et nous prouvons que $f_{\mathbb{L}}$ est injective. Par le lemme 7.35(1), nous savons qu'il existe $g\colon F \to E$ telle que $f \circ g = \mathrm{Id}\,|_F$. Nous en déduisons que $f_{\mathbb{L}} \circ g_{\mathbb{L}} = \mathrm{Id}\,|_{F_{\mathbb{L}}}$ parce que si $[\lambda, v] \in F_{\mathbb{L}}$ alors

$$(f_{\mathbb{L}} \circ g_{\mathbb{L}})[\lambda, v] = f_{\mathbb{L}}[\lambda, g(v)] = [\lambda, (f \circ g)(v)] = [\lambda, v]. \tag{9.459}$$

Notons que g est injective, donc $g_{\mathbb{L}}$ est injective et l'égalité $f_{\mathbb{L}} \circ g_{\mathbb{L}} = \mathrm{Id}\,|_{F_{\mathbb{L}}}$ implique que $f_{\mathbb{L}}$ est également injective. □

Proposition 9.204 ([1])**.**
Soit $\{e_i\}_{i=1,\dots,p}$ une base de E. Alors $\{1 \otimes e_i\}_i$ est une base de $E_{\mathbb{L}} = \mathbb{L} \otimes_{\mathbb{K}} E$.

Démonstration. L'espace vectoriel E peut être écrit comme somme directe $E = \bigoplus_i \mathbb{K}e_i$. Si $\lambda \in \mathbb{L}$ et $k \in \mathbb{K}$ nous avons

$$\lambda \otimes ke_i = \frac{\lambda}{k} \otimes e_i = \frac{\lambda}{k}(1 \otimes e_i). \tag{9.460}$$

Cela pour introduire que l'application

$$\begin{aligned} \psi\colon \mathbb{L} \otimes_{\mathbb{K}} E &\to \bigoplus_i \mathbb{L}(1 \otimes e_i) \\ \sum_k \lambda_k \otimes v_k &\mapsto \oplus_i \sum_k (\lambda_k v_{ik})(1 \otimes e_i) \end{aligned} \tag{9.461}$$

où $v_k = \sum_i v_{ik}e_i$ avec $v_{ik} \in \mathbb{K}$ est un isomorphisme de $\mathbb{L}$-espaces vectoriels. La surjectivité est facile. En ce qui concerne l'injectivité, si

$$\sum_i \sum_k (\lambda_k v_{ik})(1 \otimes e_i) = 0 \tag{9.462}$$

alors les choses suivantes sont nulles également :

$$\sum_i \sum_k (\lambda_k v_{ik})(1 \otimes e_i) = \sum_{ik} (\lambda_k \otimes v_{ik} e_i) = \sum_k (\lambda_k \otimes \sum_i v_{ik} e_i) = \sum_k (\lambda_k \otimes v_k). \tag{9.463}$$

Le dernier est l'argument de ψ. Le fait que ce soit nul implique que ψ est injective. □

Remarque 9.205.
Nous n'avons pas dû prouver que chacun des $\lambda_k \otimes v_k$ était nul. Et encore heureux, parce que cela pouvait très bien être faux, vu qu'il y a plusieurs façons de noter un élément de $E_{\mathbb{L}}$ sous la forme de tels termes.

Corollaire 9.206.
La $\mathbb{L}$-dimension de $E_{\mathbb{L}}$ est égale à la $\mathbb{K}$-dimension de E.

9.13.2 Projections

Problèmes et choses à faire

Nous allons définir $\texttt{proj}\colon \mathcal{L}(E_{\mathbb{L}}, F_{\mathbb{L}}) \to \mathcal{L}(E, F)$ en faisant appel à des bases et en prouvant que les choses définies ne dépendent pas des bases choisies. Il y a sûrement une façon plus «intrinsèque» de faire.

Nous savons que $\mathbb{L}$ est un $\mathbb{K}$-espace vectoriel dans lequel nous pouvons voir $\mathbb{K}$ comme un sous-espace (lemme 5.52). Dans cette optique nous choisissons dans $\mathbb{L}$ un supplémentaire de $\mathbb{K}$, c'est à dire un sous-espace vectoriel de $\mathbb{L}$ tel que

$$\mathbb{L} = \mathbb{K} \oplus V. \tag{9.464}$$

Nous avons alors naturellement une projection $\texttt{proj}\colon \mathbb{L} \to \mathbb{K}$.

Soit $\{e_i\}$ une base de E et $\{e_a\}$ une de$F$. Nous noterons également e_i et e_a les éléments τe_i et τe_a correspondants. Grâce à la proposition 9.204, ce sont des bases de $E_{\mathbb{L}}$ et $F_{\mathbb{L}}$. Si la fonction $f\colon E_{\mathbb{L}} \to F_{\mathbb{L}}$ s'écrit dans ce ces bases comme

$$f(e_i) = \sum_a f_{ai} e_a \tag{9.465}$$

alors nous définissons $\texttt{proj}(f)$ par

$$(\texttt{proj}\, f)e_i = \sum_a \texttt{proj}(f_{ai}) e_a. \tag{9.466}$$

Proposition 9.207 ([1])**.**
L'application $\texttt{proj}$ définie en (9.466) *est indépendante du choix des bases.*

Démonstration. Notons que dans ce qui suit, les sommes sur a ou b et celles sur i ou j ne vont pas jusqu'au même indice (dimensions de E et F). De plus nous manipulons deux choses qui se notent $\texttt{proj}$. La première est la projection $\texttt{proj}\colon \mathbb{L} \to \mathbb{K}$ qui ne dépend que d'un choix de supplémentaire et que nous supposons fixée ici. D'autre part il y a $\texttt{proj}\colon E_{\mathbb{L}} \to E$ qui dépend a priori des bases choisies.

Nous choisissons de nouvelles bases qui sont liées aux anciennes bases par

$$\begin{cases} e'_b = \sum_a B_{ab} e_a & \text{(9.467a)} \\ e'_i = \sum_j A_{ji} e_j. & \text{(9.467b)} \end{cases}$$

Les matrices A et B sont dans $\mathrm{GL}(\mathbb{K})$. Nous allons écrire l'opérateur $\texttt{proj}'$ qui correspond à ces bases et montrer que pour toute application linéaire $f\colon E_{\mathbb{L}} \to F_{\mathbb{L}}$ nous avons $\texttt{proj}(f) = \texttt{proj}'(f)$.

Nous avons :

$$f(e'_j) = \sum_i A_{ji} f(e_i) \tag{9.468a}$$
$$= \sum_a \sum_b \sum_i A_{ji} f_{ai} (B^{-1})_{ba} e'b \tag{9.468b}$$
$$= \sum_b \Big(\sum_{ai} A_{ji} f_{ai} (B^{-1})_{ba} \Big) e'b, \tag{9.468c}$$

ce qui fait que

$$(\mathtt{proj}'\, f) e'_j = \sum_b \Big(\mathtt{proj} \left(A_{ji} f_{ai} (B^{-1})_{ba} \right) \Big) e'_b. \tag{9.469}$$

Nous calculons maintenant $(\mathtt{proj}'\, f)e_j$ en substituant $e_j = \sum_l (A^{-1})_{lj} e'_l$ et en utilisant (9.469) et la linéarité de $\mathtt{proj}'$ et la $\mathbb{K}$-linéarité de $\mathtt{proj}\colon \mathbb{L} \to \mathbb{K}$:

$$(\mathtt{proj}'\, f)\Big(\sum_l (A^{-1})_{lj} e'_l \Big) = \sum_l (A^{-1})_{lj} \sum_b \sum_{ai} \mathtt{proj} \left(A_{li} f_{ai} (B^{-1})_{ba} \right) e_b \tag{9.470a}$$
$$= \sum_a \mathtt{proj}(f_{aj}) e_a \tag{9.470b}$$
$$= (\mathtt{proj}\, f) e_j. \tag{9.470c}$$

Donc $\mathtt{proj} = \mathtt{proj}'$. □

Note au passage comme toujours : il y a un abus systématique de notation entre $e_i \in E$ et $\tau(e_i) = 1 \otimes e_i \in E_{\mathbb{L}}$.

Remarque 9.208 ([1])**.**
L'opération $\mathtt{proj}\colon \mathcal{L}(E_{\mathbb{L}}, F_{\mathbb{L}}) \to \mathcal{L}(E, F)$ ne dépend pas des bases choisies un peu partout. Mais elle dépend de l'application $pr\colon \mathbb{L} \to \mathbb{K}$ déjà construite. Et celle-là dépend du choix d'un supplémentaire V qui fournit $\mathbb{L} = \mathbb{K} \oplus V$.

Si $\mathtt{proj}(\lambda) = 0$ pour un de ces choix, cela n'implique nullement que $\lambda = 0$. Penser à $i \in \mathbb{C}$ si la projection $\mathtt{proj}\colon \mathbb{C} \to \mathbb{R}$ est l'application $(x + iy) \mapsto x$ parallèle à l'axe des imaginaires.

Par contre si $\mathtt{proj}(\lambda) = 0$ pour tout choix de V, alors nous avons bien $\lambda = 0$. Dans la suit nous «fixons» un choix de V générique, et lorsque nous rencontrerons l'égalité $\mathtt{proj}(\lambda) = 0$ nous en déduirons $\lambda = 0$.

Proposition 9.209.
Si $f\colon E \to F$ et si $f_{\mathbb{L}} e_j = \sum_a (f_{\mathbb{L}})_{aj} e_a$ et si $f(e_j) = \sum_a f_{aj} e_a$ alors

(1) $\mathtt{proj}\, f_{\mathbb{L}} = f$,

(2) $(f_{\mathbb{L}})_{ja} = f_{ja} \in \mathbb{K}$.

Démonstration. Nous avons

$$f_{\mathbb{L}}(e_i) = \sum_a f_{ai} (1 \otimes e_a) = \sum_a f_{ai} \tau(e_a), \tag{9.471}$$

donc

$$(\mathtt{proj}\, f_{\mathbb{L}}) e_i = \sum_a \mathtt{proj}(f_{ai}) e_a = \sum_a f_{ai} e_a = f(e_i). \tag{9.472}$$

Cela prouve que $\mathtt{proj}\, f_{\mathbb{L}} = f$.

Par ailleurs,

$$f_{\mathbb{L}}(\tau e_i) = f_{\mathbb{L}}(1 \otimes e_i) = 1 \otimes f(e_i) = \tau\big(f(e_i)\big) = \sum_a f_{ai} \tau(e_a) \tag{9.473}$$

alors que par définition,

$$f_{\mathbb{L}}(\tau e_i) = \sum_a (f_{\mathbb{L}})_{ai} \tau(e_a). \tag{9.474}$$

Les éléments $\tau(e_a)$ formant une base[48], la comparaison de (9.473) avec (9.474) donne $(f_{\mathbb{L}})_{ai} = f_{ai} \in \mathbb{K}$. □

Lemme 9.210.
Soient

(1) Une base $\{e_i\}$ de E et une application linéaire $f\colon E \to F$;

(2) une base $\{e_a\}$ de F et une application linéaire $g\colon G \to F$;

(3) une base $\{e_\alpha\}$ de G et une application linéaire $\tilde{h}\colon G_{\mathbb{L}} \to E_{\mathbb{L}}$.

Alors nous avons

$$\texttt{proj}(f_{\mathbb{L}} \circ \tilde{h}) = \texttt{proj}(f_{\mathbb{L}}) \circ \texttt{proj}(\tilde{h}). \tag{9.475}$$

Démonstration. Pour écrire $\texttt{proj}(f_{\mathbb{L}} \circ \tilde{h})$ à partir de la définition (9.466) nous commençons par écrire

$$(f_{\mathbb{L}} \circ \tilde{h})e_\alpha = \sum_a (f_{\mathbb{L}} \circ \tilde{h})_{a\alpha} e_a = \sum_{ai} (f_{\mathbb{L}})_{ai} (\tilde{h})_{i\alpha} e_a = \sum_a \Big(\sum_i f_{ai} (\tilde{h})_{i\alpha} \Big) e_a \tag{9.476}$$

où nous avons utilisé le fait que $(f_{\mathbb{L}})_{ai} = f_{ai}$. Donc, en utilisant la $\mathbb{K}$-linéarité de $\texttt{proj}$,

$$\texttt{proj}(f_{\mathbb{L}} \circ \tilde{h})e_\alpha = \sum_a \sum_i \texttt{proj}\Big(f_{ai} (\tilde{h})_{i\alpha} \Big) e_a = \sum_a \sum_i f_{ai}\, \texttt{proj}\Big((\tilde{h})_{i\alpha} \Big) e_a. \tag{9.477}$$

D'autre part,

$$\begin{aligned} \texttt{proj}(f_{\mathbb{L}}) \circ \texttt{proj}(\tilde{h})e_\alpha &= \texttt{proj}(f_{\mathbb{L}}) \sum_i \texttt{proj}\Big((\tilde{h})_{i\alpha} \Big) e_i \\ &= \sum_i \texttt{proj}\Big((\tilde{h})_{i\alpha} \Big) \sum_a f_{ai} e_a \\ &= \sum_{ai} \texttt{proj}\Big((\tilde{h})_{i\alpha} \Big) f_{ai} e_a, \end{aligned} \tag{9.478}$$

et c'est égal à (9.477). □

Remarque 9.211.
Nous n'avons en général pas $\texttt{proj}(xy) = \texttt{proj}(x)\,\texttt{proj}(y)$ pour tout $x, y \in \mathbb{L}$. Par exemple si $\mathbb{K} = \mathbb{R}$ et $\mathbb{L} = \mathbb{C}$ avec la projection canonique,

$$\texttt{proj}(i \cdot i) = \texttt{proj}(-1) = -1 \tag{9.479}$$

alors que $\texttt{proj}(i) = 0$.

Proposition 9.212.
Soit $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, E)$. Alors il existe $h\colon G \to E$ tel que $f \circ h = g$ si et seulement s'il existe $\tilde{g}\colon G_{\mathbb{L}} \to E_{\mathbb{L}}$ tel que $f_{\mathbb{L}} \circ \tilde{g} = g_{\mathbb{L}}$.

Démonstration. Dans le sens direct, il suffit de poser $\tilde{h} = h_{\mathbb{L}}$.

Dans le sens inverse, si nous avons $\tilde{h}\colon G_{\mathbb{L}} \to E_{\mathbb{L}}$ tel que $f_{\mathbb{L}} \circ \tilde{h} = g_{\mathbb{L}}$ alors en appliquant $\texttt{proj}$ des deux côtés et en utilisant le lemme 9.210,

$$\texttt{proj}(f_{\mathbb{L}}) \circ \texttt{proj}(\tilde{h}) = \texttt{proj}(g_{\mathbb{L}}) \tag{9.480}$$

c'est à dire

$$f \circ \texttt{proj}(\tilde{h}) = g, \tag{9.481}$$

c'est à dire que l'application $\texttt{proj}\,\tilde{h}\colon G \to E$ est la réponse à la proposition. □

48. Encore la proposition 9.204.

9.13.3 Rang, polynôme minimal, polynôme caractéristique

Proposition 9.213 (Stabilité du rang par extension des scalaires[83])**.**
Si $f\colon E \to F$ est linéaire alors nous avons

$$\operatorname{rang}(f) = \operatorname{rang}(f_{\mathbb{L}}). \tag{9.482}$$

où à droite nous considérons le rang de l'application $\mathbb{L}$-linéaire $f_{\mathbb{L}}\colon E_{\mathbb{L}} \to F_{\mathbb{L}}$.

Démonstration. Il existe un supplémentaire V tel que $E = \ker(f) \oplus V$ avec $\dim(V) = \operatorname{rang}(f)$. Nous pouvons factoriser f en

$$f = f_2 \circ f_1 \tag{9.483}$$

avec $f_1\colon E \to V$ est la projection parallèle à $\ker(f)$ et est surjective (vers V) parce que $\dim(V) = \operatorname{rang}(f) = \dim\big(\operatorname{Image}(f)\big)$. De plus $f_2\colon V \to F$ est injective parce que si $v \in V$ est tel que $f_2(v) = 0$ alors on aurait

$$f(v) = (f_2 \circ f_1)(v) = f_2(v) = 0. \tag{9.484}$$

Cela donne $v \in \ker(f) \cap V = \{0\}$. Par la proposition 9.203, les applications $(f_1)_{\mathbb{L}}$ et $(f_2)_{\mathbb{L}}$ sont respectivement surjective et injective.

L'application $(f_2)_{\mathbb{L}}\colon V_{\mathbb{L}} \to F_{\mathbb{L}}$ est forcément surjective sur son image, donc

$$(f_2)_{\mathbb{L}}\colon V_{\mathbb{L}} \to \operatorname{Image}(f_{\mathbb{L}}) \tag{9.485}$$

est un isomorphisme de $\mathbb{L}$-espaces vectoriels. Nous avons alors les égalités

$$\dim_{\mathbb{L}}(V_{\mathbb{L}}) = \dim_{\mathbb{L}}\big(\operatorname{Image}(f_{\mathbb{L}})\big) = \operatorname{rang}(f_{\mathbb{L}}). \tag{9.486}$$

Mais aussi, par les définitions posées plus haut,

$$\dim(V) = \operatorname{rang}(f) = \dim\big(\operatorname{Image}(f)\big). \tag{9.487}$$

Mais le corollaire 9.206 nous dit que $\dim_{\mathbb{L}}(V_{\mathbb{L}}) = \dim_{\mathbb{K}}(V)$. Donc il y a égalité des deux lignes (9.486) et (9.487) donne $\operatorname{rang}(f) = \operatorname{rang}(f_{\mathbb{L}})$. □

Proposition 9.214.
Nous avons

(1) $\det(f) = \det(f_{\mathbb{L}})$

(2) $\chi_f = \chi_{f_{\mathbb{L}}}$.

Démonstration. Dès que l'on a des bases nous avons $(f_{\mathbb{L}})_{ai} = f_{ai}$ par la proposition 9.209(2). Le nombre $\det(f) \in \mathbb{K}$ est un polynôme en les f_{ai}. Entendons nous : il existe un polynôme indépendant de f et de $\mathbb{K}$ et de $\mathbb{L}$ donnant le déterminant de n'importe quelle matrice. Donc $\det(f) = \det(f_{\mathbb{L}})$.

Même chose pour le polynôme caractéristique (définition 9.121) : les coefficient de ce polynôme sont des polynôme en les f_{ai} qui sont indépendants de $\mathbb{L}$, de $\mathbb{K}$ et de f.

Notons que $\chi_{f_{\mathbb{L}}}$ est un polynôme à coefficients dans $\mathbb{K}$. □

La situation est très différente avec le polynôme minimal[49]. Autant il existe une «recette» pour créer le polynôme caractéristique, il n'en n'existe pas pour le polynôme minimal (ou en tout cas, il ne suffit pas d'appliquer des polynôme en les coefficients de la matrice). La proposition suivante montre que le polynôme minimal est conservé par extension de corps, mais que pour le voir, il faut travailler plus.

Proposition 9.215 ([83, 1])**.**
Soit $\mathbb{L}$ une extension du corps $\mathbb{K}$ et une application linéaire $f\colon E \to F$ entre deux $\mathbb{K}$-espaces vectoriels. Alors $\mu_f = \mu_{f_{\mathbb{L}}}$.

49. Définition 5.48.

Démonstration. Nous allons montrer que l'application

$$\begin{aligned}\tilde{g}\colon \frac{\mathbb{L}[X]}{(\mu)} &\to \text{End}(E_{\mathbb{L}})\\ \bar{P} &\mapsto P(f_{\mathbb{L}})\end{aligned} \tag{9.488}$$

est bien définie et injective. La proposition 9.111 nous dira alors que μ est le polynôme minimal de $f_{\mathbb{L}}$.

Pour prouver que l'application $\tilde{g}$ est bien définie, nous commençons par prouver que $P(f_{\mathbb{L}}) = P(f)_{\mathbb{L}}$:

$$P(f_{\mathbb{L}})\lambda \otimes v = \sum_k a_k f_{\mathbb{L}}^k \lambda \otimes v \tag{9.489a}$$
$$= \lambda \otimes \sum_k a_k f^k(v) \tag{9.489b}$$
$$= \lambda \otimes P(f)v \tag{9.489c}$$
$$= P(f)_{\mathbb{L}}\lambda \otimes v. \tag{9.489d}$$

Par conséquent $\mu(f_{\mathbb{L}}) = 0$ et l'application est bien définie.

Sur $\mathbb{L}[X]/(\mu)$ nous considérons la base $\{1, \bar{X}, \dots, \bar{X}^{\deg(\mu)-1}\}$, et $\text{End}(E_{\mathbb{L}})$ nous considérons une base qui commence [50] par $\{f_{\mathbb{L}}^k\}_{k=0,\dots,\deg(\mu)-1}$. Montrons tout de même que cette partie est libre (sinon le théorème de la base incomplète ne s'applique pas) : si $\sum_k \lambda_k f_{\mathbb{L}}^k = 0$ alors

$$\sum_k \texttt{proj}\left(\lambda_k f_{\mathbb{L}}^k\right) = 0. \tag{9.490}$$

Pour détailler ce que cela implique, nous calculons ceci :

$$(\lambda f_{\mathbb{L}})(\tau e_i) = \lambda f_{\mathbb{L}}(\tau e_i) = \sum_a \lambda f_{ia} e_a, \tag{9.491}$$

par conséquent $\texttt{proj}(\lambda f_{\mathbb{L}})e_i = \sum_a \texttt{proj}(\lambda f_{ia})e_a$, et comme $\texttt{proj}$ est $\mathbb{K}$-linéaire et que $f_{ai} \in \mathbb{K}$,

$$\texttt{proj}(\lambda f_{\mathbb{L}})e_i = \texttt{proj}(\lambda)\sum_a f_{ai}e_a = \texttt{proj}(\lambda)\,\texttt{proj}(f_{\mathbb{L}})e_i = \texttt{proj}(\lambda)f(e_i). \tag{9.492}$$

Appliquer la projection $\texttt{proj}$ à l'équation (9.490) donne alors $\sum_k \texttt{proj}(\lambda)_k f^k = 0$. Mais comme les f^k sont linéairement indépendantes sur $\mathbb{K}$ nous avons pour tout k : $\texttt{proj}(\lambda_k) = 0$ (égalité dans $\mathbb{K}$). En nous souvenant de la remarque 9.208 nous en déduisons $\lambda_k = 0$ dans $\mathbb{L}$.

Dans les choix de bases faits, l'application $\tilde{g}$ a la forme

$$\tilde{g} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ * & * & * & \\ * & * & * & \\ * & * & * & \end{pmatrix}, \tag{9.493}$$

qui est injective.

Vu que $\tilde{g}$ est injective, μ est le polynôme minimal de $f_{\mathbb{L}}$ et donc $\mu = \mu_{\mathbb{L}}$. □

9.14 Frobenius et Jordan

9.14.1 Matrice compagnon

Définition 9.216.
*Soit le polynôme $P = X^n - a_{n-1}X^{n-1} - \dots - a_1X - a_0$ dans $\mathbb{K}[X]$. La **matrice compagnon** de*

50. Théorème de la base incomplète 7.10(2).

P est la matrice donnée par

$$C(P) = \begin{pmatrix} 0 & \cdots & \cdots & 0 & a_0 \\ 1 & 0 & & \vdots & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_{n-2} \\ 0 & \cdots & 0 & 1 & a_{n-1} \end{pmatrix} \tag{9.494}$$

si $n \geqslant 2$ et par (a_0) si $n = 1$.

Une matrice est dite compagnon si elle a cette forme.

Proposition 9.217.
Si f est l'endomorphisme associé à la matrice $C(P)$ nous avons

$$f(e_i) = \begin{cases} e_{i+1} & si\ i < n \\ (a_0, \ldots, a_{n-1}) & si\ i = n. \end{cases} \tag{9.495}$$

De plus l'endomorphisme f vérifie $P(f)e_1 = 0$.

Lemme 9.218 ([110])**.**
Un polynôme sur un corps commutatif est le polynôme caractéristique de sa matrice compagnon. En d'autres termes nous avons $\chi_{C(P)} = P$.

Démonstration. Nous notons f l'endomorphisme associé à $C(P)$. La propriété $P(f)e_1 = 0$ nous indique que le polynôme minimal ponctuel de f en e_1 divise P. L'ensemble des puissances de f appliquées à e_1, $\big(f^i(e_1)\big)_{i=1,\ldots,n-1}$ est libre, donc le polynôme minimal ponctuel en e_1 est de degré n au minimum. En reprenant les notations du théorème 5.37, nous avons $I_{e_1} = (P)$ parce que P est de degré minimum dans I_{e_1} et $\chi_f \in I_{e_1}$.

Donc P divise χ_f et est de degré égal à celui de χ_f. Étant donné qu'ils sont tous deux unitaires, ils sont égaux. □

Remarque 9.219.
Les matrices compagnons ne sont pas les seules dont le polynôme caractéristique est égal au polynôme minimal. En fait les matrices dont le polynôme caractéristique est égale au polynôme minimal sont denses dans les matrices. En effet une matrice dont le polynôme minimal n'est pas égal au polynôme caractéristique a un polynôme caractéristique avec une racine double. Il est possible, en modifiant arbitrairement peu la matrice de séparer la racine double en deux racines distinctes.

9.14.2 Réduction de Frobenius

Lemme 9.220.
Soit un endomorphisme $f : E \to E$ sur l'espace vectoriel de dimension finie n. Nous notons μ et χ les polynômes minimal et caractéristique. Si f est cyclique, alors $\mu = \chi$.

Le théorème 9.233 donnera une version plus complète de ce lemme.

Démonstration. Soit v un vecteur cyclique de f, c'est à dire que $\{f^k(v)\}_{k=0,\ldots,n-1}$ est libre. Donc si P est un polynôme de degré jusqu'à $n-1$ nous ne pouvons pas avoir $P(f) = 0$ parce que, appliqué à v, ce serait une combinaisons nulle non triviale des $f^k(v)$. Donc le polynôme minimal est au minimum de degré n. Mais le polynôme caractéristique est annulateur de degré n (Cayley-Hamilton 9.130), donc il est le polynôme minimal. □

Théorème 9.221 (Réduction de Frobenius [111, 112, 113])**.**
Soit E, un $\mathbb{K}$-espace vectoriel, et $f \in \mathrm{End}(E)$. Alors il existe une suite de sous-espaces $E_1, \ldots, E_r$ stables par f tels que

(1) $E = \bigoplus_{i=1}^{r} E_i$;

(2) pour chaque E_i*, l'endomorphisme restreint* $f_i = f|_{E_i}$ *est cyclique ;*

(3) si μ_i *est le polynôme minimal de* f_i *alors* μ_{i+1} *divise* μ_i *;*

Une telle décomposition vérifie automatiquement $\mu_1 = \mu_f$ *et* $\mu_1 \cdots \mu_r = \chi_f$*, et la suite* $(\mu_i)_{i=1,\dots,r}$ *ne dépend que de* f *et non du choix de la décomposition du point (1).*

Les polynômes μ_i sont les **invariants de similitude** de l'endomorphisme f.

Démonstration. Nous commençons par montrer que si une telle décomposition existe, alors

$$\chi_f = \prod_{i=1}^{r} \mu_i \tag{9.496a}$$

$$\mu_f = \mu_1 \tag{9.496b}$$

où χ_f est le polynôme caractéristique de f et μ_f est le polynôme minimal. D'abord le polynôme caractéristique de f devra être égal au produit des polynômes caractéristique des $f|_{E_i}$, mais ces derniers endomorphismes étant cycliques[51], leurs polynôme caractéristiques sont égaux à leurs polynômes minimaux (lemme 9.220). Cela prouve l'égalité (9.496a). Ensuite tous les μ_i doivent diviser le polynôme minimal, donc $\operatorname{ppcm}(\mu_1, \dots, \mu_r)$ divise μ_f. Cependant le polynôme minimal doit contenir une et une seule fois chacun des facteurs irréductibles du polynôme caractéristique, et chacun de ces facteurs sont dans les polynômes μ_i. Par conséquent $\operatorname{ppcm}(\mu_1, \dots, \mu_r) = \mu_f$. Mais par ailleurs $\mu_1 = \operatorname{ppcm}(\mu_1, \dots, \mu_r)$ parce qu'on a supposé $\mu_{i+1} \mid \mu_i$, donc $\mu_1 = \mu_f$.

Soit d, le degré du polynôme minimal de f et $y \in E$ tel que $\mu_f = \mu_{f,y}$ (voir lemme 9.112). Le plus petit espace stable sous f contenant y est

$$E_y = \operatorname{Span}\{y, f(y), \dots, f^{d-1}(y)\}. \tag{9.497}$$

Nous notons $e_i = f^{i-1}(y)$. Notons que les vecteurs donnés forment bien une base de E_y parce que si les e_i n'était pas linéairement indépendants, alors nous aurions des a_k tels que $\sum_k a_k e_k = 0$ et avec lesquels

$$\big(\sum_k a_k X^k\big)(f)y = 0, \tag{9.498}$$

ce qui contredirait la minimalité de $\mu_{f,y}$.

La difficulté du théorème est de trouver un complément de E_y qui soit également stable sous f. Nous commençons par étendre[52] $\{e_1, \dots, e_d\}$ en une base $\{e_1, \dots, e_n\}$ de E. Ensuite nous allons montrer que

$$E = E_y \oplus F \tag{9.499}$$

avec

$$F = \{x \in E \text{ tel que } e_d^*\big(f^k(x)\big) = 0 \forall k \in \mathbb{N}\}. \tag{9.500}$$

Par construction, F est invariant sous f. Montrons pour commencer que $E_y \cap F = \{0\}$. Un élément de E_y s'écrit

$$z = a_1 e_1 + \cdots + a_k e_k \tag{9.501}$$

avec $k \leqslant d$. Étant donné que f décale les vecteurs de base, nous avons $e_d^*\big(f^{d-k}(z)\big) = a_k$. Du coup $z \in F$ si et seulement si $a_1 = \ldots = a_d = 0$, c'est à dire que $E_y \cap F = \{0\}$.

Nous montrons maintenant que $\dim F = n - d$. Pour cela nous considérons l'application

$$\begin{aligned} T\colon \mathbb{K}[F] &\to E^* \\ g &\mapsto e_d^* \circ g. \end{aligned} \tag{9.502}$$

Cette application est injective. En effet un élément général de $\mathbb{K}[f]$ est

$$g = a_1 \operatorname{Id} + a_2 f + \cdots + a_p f^{p-1} \tag{9.503}$$

51. Définition 9.113.

52. Pour autant que j'aie compris, cette extension manque dans [111]. Corrigez moi si je me trompe.

avec $p \leqslant d$. Si $T(g) = 0$, alors nous avons en particulier

$$0 = T(g)e_{d-p+1} = e_d^*(a_1e_{d-p+1} + a_2e_{d-p+2} + \cdots + a_pe_d) = a_p. \tag{9.504}$$

Donc $a_p = 0$ et en appliquant maintenant $T(g)$ à e_{d-p} nous obtenons $a_{p-1} = 0$. Au final nous trouvons que $g = 0$ et donc que T est injective.

Étant donné que $\dim \mathbb{K}[f] = d$ et que T est injective, $\dim \text{Image}(T) = d$. Nous regardons l'orthogonal de l'image :

$$(\text{Image}(T))^\perp = \{x \in E \text{ tel que } T(g)x = 0 \forall g \in \mathbb{K}[f]\} \tag{9.505a}$$
$$= \{x \in E \text{ tel que } e_d^*\big(g(x)\big) = 0 \forall g \in \mathbb{K}[f]\} \tag{9.505b}$$
$$= F. \tag{9.505c}$$

Par conséquent $F^\perp = \text{Image}(T)$. Vu que $\dim \text{Image}(T) = d$, nous avons donc $\dim F = n - d$ et il est établi que $E = E_y \oplus F$.

Nous avons donc trouvé F, stable par f et tel que $E = E_y \oplus F$. Nous devons maintenant nous assurer que cette décomposition tombe bien pour les polynômes minimaux. Si P_1 est le polynôme minimal de $f|_{E_yj}$, alors par le lemme 9.114 nous avons $P_1 = \mu_{f,y} = \mu_f$ parce que $f|_{E_y}$ est cyclique sur E_y. Mettons P_2, le polynôme minimal de $f|_F$. Étant attendu que F est stable par f, le polynôme P_2 divise P_1. En recommençant la construction sur F, nous construisons un nouvel espace F' stable sous F et vérifiant $\mu_{f|_{F'}} = P_2$, etc.

Nous passons maintenant à la partie unicité du théorème. Soient deux suites $F_1, \ldots, F_r$ et $G_1, \ldots, G_s$ de sous-espaces stables par f et vérifiant

(1) $E = \bigoplus_{i=1}^r F_i$,

(2) $f|_{F_i}$ est cyclique,

(3) $\mu_{f|_{F_{i+1}}}$ divise $\mu_{f|_{F_i}}$,

et, *mutatis mutandis*, les mêmes conditions pour la famille $\{G_i\}$. Nous posons $P_i = \mu_{f_{F_i}}$ et $Q_i = \mu_{f|_{G_i}}$. Nous allons montrer par récurrence que $P_i = Q_i$ et $\dim F_i = \dim G_i$. Il ne sera cependant pas garanti que $F_i = G_i$. D'abord, $P_1 = Q_1$ parce qu'ils sont tous deux égaux à μ_f par les relations (9.496). Nous supposons que $P_i = Q_i$ pour $i \leqslant 1 \leqslant j - 1$ et nous tentons de montrer que $P_j = Q_j$.

Nous avons

$$P_j(f) = P_j(f)|_{F_1} \oplus \ldots \oplus P_j(f)|_{F_{j-1}}. \tag{9.506}$$

En effet étant donné que P_{j+k} divise P_j, nous avons[53] $P_j(f) = A(f) \circ P_{j+k}(f)$, mais $P_{j+k}(f)F_{j+k} = 0$, donc $P_j(f)F_{j+k} = 0$. Les espaces G_i n'ayant a priori aucun rapport avec les polynômes P_i, nous écrivons

$$P_j(f) = P_j(f)|_{G_1} \oplus \ldots \oplus P_j(f)|_{G_{j-1}} \oplus P_j(f)|_{G_j} \oplus \ldots \oplus P_j(f)|_{G_s}. \tag{9.507}$$

Pour $1 \leqslant i \leqslant j - 1$, nous avons supposé $P_i = Q_i$. Étant donné que $f|_{F_i}$ est semblable à C_i et $f|_{G_i}$ est semblable à C_{Q_i}, la matrice de $f|_{E_i}$ est semblable à la matrice de $f|_{G_i}$. En particulier,

$$\dim P_j(f)F_i = \dim P_j(f)G_i. \tag{9.508}$$

En prenant les dimensions des images dans les égalités (9.506) et (9.507), nous trouvons que

$$P_j(f)|_{G_j} = \ldots = P_j(f)|_{G_s} = 0. \tag{9.509}$$

Par conséquent $P_j \in I_{f|G_j}$ et donc P_j divise Q_j, qui est générateur de $I_{f|_{G_j}}$. La situation étant symétrique entre P et Q, nous montrons de même que Q_j divise P_j et donc que $P_j = Q_j$.

Ceci achève la démonstration du théorème de réduction de Frobenius.

□

53. En vertu du lemme 9.101.

Remarque 9.222.
Sous forme matricielle, ce théorème dit que toute matrice est semblable à une matrice de la forme bloc-diagonale

$$f = \begin{pmatrix} C_{\mu_1} & & \\ & \ddots & \\ & & C_{\mu_r} \end{pmatrix} \tag{9.510}$$

où les C_{μ_i} sont les matrices compagnon (définition 9.216).

En particulier, et ceci est très important, deux applications sont semblables si et seulement si elles ont même suite d'invariants de similitude.

Remarque 9.223.
Si nous travaillons sur $\mathbb{R}$, la réduite de Frobenius restera une matrice réelle, même si les valeurs propres sont complexes. En effet le procédé de Frobenius ne regarde absolument pas les valeurs propres, mais seulement les facteurs irréductibles du polynôme caractéristique. La réduite de Frobenius ne tente pas de résoudre ces polynômes, mais se contente d'en utiliser les matrices compagnon.

La situation sera différente dans le cas de la forme normale de Jordan.

9.14.3 Forme normale de Jordan

Il existe une preuve directe de la réduction de Jordan ne nécessitant pas la réduction de Frobenius[105]. Cette dernière passe par les espaces caractéristiques [54] et est à mon avis plus compliquée que la démonstration de Frobenius elle-même. Nous allons donc nous contenter de donner la réduction de Jordan comme un cas particulier de Frobenius.

Théorème 9.224 (Réduction de Jordan)**.**
Soit E un espace vectoriel sur $\mathbb{K}$, et $f \in \mathrm{End}(E)$ un endomorphisme dont le polynôme caractéristique χ_f est scindé[55]*. Il existe une base de E dans laquelle la matrice de f s'écrit sous la forme*

$$M = \begin{pmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_k}(\lambda_k) \end{pmatrix} \tag{9.511}$$

où les λ_i sont les valeurs propres de f (avec éventuelle répétitions) et $J_n(\lambda)$ représente le bloc $n \times n$

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \lambda & & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}. \tag{9.512}$$

En d'autres termes, $J_n(\lambda)_{ii} = \lambda$ et $J_n(\lambda)_{i-1,i} = 1$.

Démonstration. Nous commençons par le cas où f est nilpotente ; nous notons M sa matrice. Dans ce cas la seule valeur propre est zéro et le polynôme caractéristique est X^m pour un certain m. Nous savons par le lemme 9.218 que (la matrice de) f est semblable à sa matrice compagnon. En l'occurrence pour f nous avons

$$C_{X^m} = \begin{pmatrix} 0 & & & 0 \\ 1 & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix}. \tag{9.513}$$

54. Aussi appelés «espaces propres généralisés».
55. C'est pour cette hypothèse que $\mathbb{K} = \mathbb{R}$ n'est pas le bon cadre.

Ensuite le changement de base (qui est une similitude) $(e_1, \dots, e_n) \mapsto (e_n, \dots, e_1)$ montre que C_{X^m} est semblable à un bloc de Jordan $J_m(0)$.

Supposons à présent que f ne soit pas nilpotente. Par l'hypothèse de polynôme caractéristique scindé, nous supposons que f a m valeurs propres distinctes et que son polynôme caractéristique est

$$\chi_f = (X - \lambda_1)^{l_1} \dots (X - \lambda_m)^{l_m}. \tag{9.514}$$

Le lemme des noyaux (théorème 9.102) nous enseigne que

$$E = \bigoplus_{i=1}^{m} \underbrace{\ker(f - \mu_i \mathbb{1})^{l_i}}_{F_i}. \tag{9.515}$$

La restriction de $f - \lambda_i \mathbb{1}$ à F_i est par construction un endomorphisme nilpotent, et donc peut s'écrire comme un bloc de Jordan avec des zéros sur la diagonale. En utilisant la décomposition

$$f|_{F_i} = (f - \lambda_i \mathbb{1})|_{F_i} + \lambda_i \mathbb{1}_{F_i}, \tag{9.516}$$

nous voyons que $f|_{F_i}$ s'écrit comme un bloc de Jordan avec λ_i sur la diagonale. □

Remarque 9.225.
Nous pouvons calculer la forme normale de Jordan pour une matrice complexe ou réelle, mais dans les deux cas nous devons nous attendre à obtenir une matrice complexe parce que les valeurs propres d'une matrice réelle peuvent être complexes. Cependant nous demandons que le polynôme caractéristique de f soit scindé sur $\mathbb{K}$. En pratique, la décomposition de Jordan n'est garantie que sur les corps algébriquement clos, c'est à dire sur $\mathbb{C}$.

La suite des invariants de similitude sur laquelle repose Frobenius, elle, est disponible sur tout corps, y compris $\mathbb{R}$.

9.15 Commutant et endomorphismes cycliques

9.15.1 Endomorphisme cyclique

Lemme 9.226.
Si A est la matrice de l'endomorphisme f alors nous avons équivalence des propriétés suivantes :

(1) La matrice A est cyclique.

(2) L'endomorphisme f est cyclique.

Si f est un endomorphisme de l'espace vectoriel E et si $x \in E$, nous notons

$$E_{f,x} = \text{Span}\{f^k(x) \text{ tel que } k \in \mathbb{N}\}. \tag{9.517}$$

Définition 9.227.
*Soit E un espace vectoriel de dimension finie sur un corps $\mathbb{K}$ et un endomorphisme $f : E \to E$. Le **commutant** de f est l'ensemble des endomorphismes de E qui commutent avec f :*

$$\mathcal{C}(f) = \{g \in \mathcal{L}(E, E) \text{ tel que } g \circ f = f \circ g\}. \tag{9.518}$$

Il n'est pas très compliqué de vérifier que $\mathcal{C}(f)$ est un sous-espace vectoriel de $\mathcal{L}(E, E)$.

Notons l'inclusion évidente $\mathbb{K}[f] \subset \mathcal{C}(f)$. L'inclusion inverse va un peu nous occuper durant les prochaines pages.

9.15.2 Commutant : cas diagonalisable

Proposition 9.228 ([114]).
Si f est diagonalisable, alors

$$\dim\big(\mathcal{C}(f)\big) = \sum_{\lambda\in\mathrm{Spec}(f)} \dim(E_\lambda)^2. \tag{9.519}$$

où les E_λ sont les espaces propres de f.

Démonstration. D'abord su $g \in \mathcal{C}(f)$ alors E_λ est stable par g. En effet si $v \in E_\lambda$ alors $f\big(g(v)\big) = g\big(f(v)\big) = g(\lambda v) = \lambda g(v)$, ce qui montre que $g(v)$ est un vecteur propre de f pour la valeur propre λ, et donc que $g(v) \in E_\lambda$.

Nous considérons ensuite l'application

$$\begin{aligned}\psi\colon \mathcal{C}(f) &\to \mathrm{End}(E_1) \times \ldots \times \mathrm{End}(E_r)\\ g &\mapsto g|_{E_1} \times \ldots \times g|_{E_r}\end{aligned} \tag{9.520}$$

qui est bien définie parce que g se restreint aux espaces propres de f. Nous allons noter $\psi(g)_\lambda$ la restriction de g à E_λ.

ψ **est injective** Supposons que $g, h \in \mathcal{C}(f)$ tels que $\psi(g) = \psi(h)$. Vu que f est diagonalisable nous pouvons décomposer $x \in E$ en ses composantes sur les espaces propres [56] :

$$x = \sum_{\lambda\in\mathrm{Spec}(f)} x_\lambda \tag{9.521}$$

avec $x_\lambda \in E_\lambda$. Nous avons alors

$$g(x) = \sum_\lambda g(x_\lambda) = \sum_\lambda \psi(g)_\lambda(x_\lambda). \tag{9.522}$$

Vu que nous avons $\psi(g)_\lambda = \psi(h)_\lambda$, nous avons aussi

$$g(x) = \sum_\lambda \psi(g)_\lambda(x_\lambda) = \sum_\lambda \psi(h)_\lambda(x_\lambda) = \sum_\lambda h(x_\lambda) = h(x). \tag{9.523}$$

Cela prouve $g = h$ et donc que ψ est injective.

ψ **est surjective** Si nous avons pour chaque $\lambda \in \mathrm{Spec}(f)$ un endomorphisme g_λ de E_λ alors en posant

$$g(x) = \sum_{\lambda\in\mathrm{Spec}(f)} g_\lambda(x_\lambda) \tag{9.524}$$

alors nous avons bien

$$\psi(g) = \big(g_{\lambda_1}, \ldots, g_{\lambda_r}\big). \tag{9.525}$$

Nous pouvons donc conclure en écrivant

$$\dim\big(\mathcal{C}(f)\big) = \sum_{\lambda\in\mathrm{Spec}(f)} \dim\big(\mathrm{End}(E_\lambda)\big) = \sum_{\lambda\in\mathrm{Spec}(f)} \dim(E_\lambda)^2. \tag{9.526}$$

□

Remarque 9.229.
Nous avons alors immédiatement

$$\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E) \tag{9.527}$$

lorsque f est diagonalisable.

56. Théorème 9.143(5).

En suivant la notation (9.243), un endomorphisme est cyclique lorsqu'il existe $x \in E$ tel que $E_x = E$.

Proposition 9.230 ([114])**.**
Si f est un endomorphisme diagonalisable d'un espace vectoriel E de dimension n. Nous avons équivalence entre les faits suivants.

(1) Le polynôme minimal est égal au polynôme caractéristique : $\mu_f = \chi_f$
(2) L'endomorphisme f est cyclique.
(3) $\mathcal{C}(f) = \mathbb{K}[f]$.
(4) $\dim\big(\mathcal{C}(f)\big) = n$
(5) L'endomorphisme f possède n valeurs propres distinctes.
(6) $\dim\big(\mathbb{K}[f]\big) = n$

Démonstration. Le point important de cette proposition sont les équivalences (1)-(3). Les autres sont des intermédiaires. En particulier, dans le cas diagonalisable, nous allons voir que le point (5) est essentiellement une reformulation de (1).

(4) implique (5) Par la formule (9.519), les espaces propres de f ont dimension 1. Par conséquent f possède n valeurs propres distinctes.

(5) implique (6) Le théorème 9.143 nous dit que le polynôme minimal est scindé à racines simples. Vu que f possède n valeurs propres distinctes, μ est de degré n. Par l'isomorphisme $\mathbb{K}[f] = \mathbb{K}[X]/(\mu)$ de la proposition 9.120 nous avons $\dim\big(\mathbb{K}[f]\big) = \deg(\mu) = n$ par la proposition 5.38.

(6) implique (1) Par l'isomorphisme $\mathbb{K}[f] = \mathbb{K}[X]/(\mu)$ de la proposition 9.120 et la proposition 5.38 nous avons $n = \dim\big(\mathbb{K}[f]\big) = \deg(\mu)$. Vu que χ est un polynôme annulateur (Caley-Hamilton 9.130), il est divisé par μ. Maintenant μ et χ sont des polynômes unitaires de degré n et μ divise χ. Ils sont donc égaux.

(1) implique (2) Le fait que f soit diagonalisable permet d'utiliser le théorème 9.143 pour dire que μ est scindé à racines simples. L'égalisation avec χ nous permet de dire que f possède n valeurs propres distinctes. Soient $\{e_1, \dots, e_n\}$ une base de diagonalisation, et prouvons que le vecteur $v = e_1 + \cdots + e_n$ est cyclique. Nous avons

$$f^k(v) = \sum_{i=1}^{n} \lambda_i^k e_i. \tag{9.528}$$

Pour prouver que cette famille (avec $k = 0, \dots, n-1$) est libre[57] nous en prenons une combinaison linéaire nulle et nous prouvons que les coefficients sont tous nuls. Soit donc

$$0 = \sum_{l=0}^{n-1} a_l f^l(v) = \sum_{l=0}^{n-1} a_l \sum_{i=1}^{n} \lambda_i^l e_i = \sum_{i=1}^{n} \Big(\sum_{l=0}^{n-1} a_l \lambda_i^l \Big) e_i. \tag{9.529}$$

Vu que cela est nul, nous avons pour tout i :

$$\sum_{l=0}^{n-1} a_l \lambda_i^l = 0. \tag{9.530}$$

En posant la matrice $A_{ij} = \lambda_i^j$, cela revient à étudier le système $\sum_j A_{ij} a_j = 0$. Ce système n'a des solutions non nulles que si $\det(A) = 0$; sinon il possède une unique solution et elle est $a_j = 0$ pour tout j. Nous devons donc calculer le déterminant

$$\det \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{pmatrix}. \tag{9.531}$$

57. Ce sera alors une base parce que n vecteurs libres dans un espace de dimension n est toujours une base, théorème 7.14(3).

Il s'agit du déterminant de Vandermonde déjà étudié par la proposition 9.12. Nous avons $\det(A) = \prod_{1\leqslant i<j\leqslant n}(\lambda_j - \lambda_i)$. Cela est bien non nul du fait que toutes les valeurs propres soient distinctes.

(2) implique (3) Soit v un vecteur cyclique de f. Un endomorphisme g donne lieu à un polynôme par le fait suivant : il existe des uniques a_k $(k = 0, \dots, n-1)$ tels que

$$g(v) = \sum_{k=0}^{n-1} a_k f^k(v). \tag{9.532}$$

Cela donne une application linéaire

$$\begin{aligned}\psi\colon \mathcal{C}(f) &\to \mathbb{K}[f]\\ g &\mapsto P \text{ tel que } P(f)v = g(v).\end{aligned} \tag{9.533}$$

C'est une application injective parce que si $\psi(g) = 0$ alors $g(v) = 0$ et pour tout k nous avons $g\big(f^k(v)\big) = f^k\big(g(v)\big) = 0$. L'endomorphisme g s'annulant sur une base, est nul.

(3) implique (4) Si $n_1, \dots, n_r$ sont les dimensions des différents espaces propres de f, nous avons les inégalités

$$\dim\big(\mathbb{K}[f]\big) = \deg(\mu) \leqslant n = n_1 + \cdots + n_r \leqslant n_1^2 + \cdots + n_r^2 = \dim\big(\mathcal{C}(f)\big). \tag{9.534}$$

Par hypothèse d'égalité entre le premier et le dernier terme de cette suite d'inégalités, toutes les inégalités sont des égalités et en particulier $\dim\big(\mathcal{C}(f)\big) = n$.

Nous avons fini de prouver toutes les équivalences demandées. □

Exemple 9.231
Pour mieux comprendre pourquoi le fait d'avoir n valeurs propres distinctes est équivalent à être cyclique, notons que si deux valeurs propres sont identiques, alors un morceau de la matrice de f serait par exemple $\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}$, et dans ce cas n'importe quelle combinaison $ae_i + be_j$ reste proportionnelle à elle-même après application de f. Si nous avons des valeurs propres différentes par contre, nous avons par exemples dans $\mathbb{R}^2$ la matrice $\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$ qui donne $f(e_1 + e_2) = e_1 + 2e_2$. La partie $\{e_1 + e_2, e_1 + 2e_2\}$ est une base. △

9.15.3 Commutant : cas général

Nous considérons encore un espace vectoriel E de dimension finie n et un endomorphisme $f\colon E \to E$. Nous notons μ sont polynôme minimal et μ_x le polynôme minimal ponctuel en x.

Lemme 9.232 ([115, 111, 83])**.**
Nous avons

$$\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E) \tag{9.535}$$

Démonstration. Si f est donnée, l'espace $\mathcal{C}(f)$ est l'espace des solutions de $fg = gf$. Supposons avoir choisit une base de E et notons A la matrice de f et X celle de g. L'équation est $AX - XA = 0$.

Si A est trigonalisable Nous supposons avoir choisi la base de telle sorte que A soit triangulaire supérieure, et nous allons nous contenter de chercher les solutions X qui sont également triangulaires supérieure. S'il y en a déjà plus que n, a fortiori le résultat sera vrai.

Le produit de deux matrices triangulaires supérieures étant une matrice triangulaire supérieure, l'équation $AX - XA$ contient, pour les coefficients de X, $n(n+1)/2$ équations. Mais il se fait que les termes diagonaux ne sont pas de vraies équations parce que

$$(AX - XA)_{kk} = \sum_i \big(A_{ki}X_{ik} - X_{ki}A_{ik}\big) = \sum_{k\leqslant i\leqslant k} (A_{ki}X_{ik} - X_{ki}A_{ik}) = 0. \tag{9.536}$$

Nous avons donc au maximum

$$\frac{n(x+1)}{2} - n \tag{9.537}$$

équations linéairement indépendantes pour un minimum de $n(n+1)/2$ inconnues. L'espace des solutions est donc de dimension au minimum n.

Cela a l'air d'être une majoration assez large, mais il existe des cas d'égalité.

Si A n'est pas trigonalisable La preuve que nous donnons ici est valable même pour les endomorphismes trigonalisables.

Nous considérons le résultat de Frobenius 9.221. Nous avons donc la structure suivante :
— une décomposition en somme directe $E = E_1 \oplus \dots \oplus E_r$,
— les espaces E_i sont fixés par f,
— les endomorphismes $f_i = f|_{E_i}$ sont cycliques
— le polynôme minimal de f_i est μ_i et $\prod_{i=1}^{r} \mu_i = \chi_f$.
Les endomorphismes f_i^k commutent évidemment avec f_j, et la partie $\{f_i^k\}_{k=0,\dots,\deg(\mu_i)-1}$ est libre. Libre en tout cas en tant que partie de $\mathrm{End}(E_i)$. Mais en prolongeant par 0 sur E, ça reste libre en tant que partie de $\mathrm{End}(E)$.

Bien entendu les f_j^k et les f_i^k $(i \neq j)$ sont linéairement indépendants dans $\mathrm{End}(E)$ parce qu'ils n'agissent pas sur les mêmes vecteurs. Donc les endomorphismes $f_i^{k_i}$ avec $k_i = 0, \dots, \deg(\mu_i) - 1$ forment une partie libre de $\mathrm{End}(E)$ composée d'endomorphismes qui commutent avec f. Il y en a en tout

$$\sum_{i=1}^{r} \deg(\mu_i) = \deg(\chi_f) = n. \tag{9.538}$$

Par conséquent $\dim\big(\mathcal{C}(f)\big) \geqslant \dim(E)$.

□

Théorème 9.233 ([99])**.**
Soit un endomorphisme $f\colon E \to E$ sur l'espace vectoriel de dimension finie n. Nous notons μ et χ les polynômes minimal et caractéristique. Nous avons équivalence entre les faits suivants :

(1) $\mu = \chi$,
(2) f *est cyclique,*
(3) $\mathcal{C}(f) = \mathbb{K}[f]$.

Démonstration. Plusieurs implications. Notons que (1) implique (1) a déjà été démontré par le lemme 9.220.

(1) implique (2) Conformément à ce que nous permet le lemme 9.112 nous choisissons[58] $a \in E$ de telle sorte à avoir $\mu_a = \mu$. De plus pour $x \in E$ nous considérons l'application

$$\begin{aligned} \varphi_x\colon \mathbb{K}[X] &\to E \\ P &\mapsto P(f)x. \end{aligned} \tag{9.539}$$

Nous avons $\varphi_a(P) = P(f)a$ et vu que E_a est engendré par les $f^k(a)$ nous avons $\varphi_a\big(\mathbb{K}[X]\big) = E_a$. De plus l'application φ_a passe aux classes pour (μ_a). Pour rappel, un élément de $\mathbb{K}[X]/(\mu_a)$ est de la forme

$$\bar{P} = \{P + Q\mu_a\}_{Q \in \mathbb{K}[X]}. \tag{9.540}$$

Nous considérons donc l'application

$$\varphi_a\colon \frac{\mathbb{K}[X]}{(\mu_a)} \to E_a \tag{9.541}$$

et nous prouvons que c'est un isomorphisme d'espace vectoriel.

58. Dans toute la suite,nous devrions écrire μ_f et $\mu_{f,a}$ mais nous omettons d'indiquer explicitement la dépendance en f.

Linéaire Parce que $(\lambda P + Q)(f) = (\lambda P)(f) + Q(f)$.

Injectif Si $\varphi_a(\bar{P}) = 0$ alors $\varphi_a(P) = 0$ (dans la deuxième, φ_a est l'application définie sur les polynômes et non sur les classes), ce qui montrer que P est annulateur de a. Mais par définition 9.108 du polynôme minimal ponctuel, μ_a est générateur de $\ker(\varphi_a)$; donc il existe $Q \in \mathbb{K}[X]$ tel que $P = Q\mu_a$. En d'autres termes, du point de vue du quotient, $\bar{P} = 0$.

Surjectif Si $x \in E_a$ alors il existe des coefficients $x_k \in \mathbb{K}$ tels que $x = \sum_{k=0}^{\deg(\mu_a)-1} x_k f^k(a)$, c'est à dire $x = P(f)a = \varphi_a(P)$.

Mais par hypothèse et par choix de a nous avons $\mu_a = \mu = \chi$, donc en fait $E_a = \mathbb{K}[X]/(\chi)$. Mais nous savons que $\deg(\chi) = \dim(E)$ et que $\dim\big(\mathbb{K}[X]/P\big) = \deg(P)$ par la proposition 9.120. Au final nous avons $\dim(E_a) = \deg(\chi) = \dim(E)$. Et par conséquent $E_a = E$. Cela prouver que a est un vecteur cyclique pour f.

(2) implique (3) Soit $g \in \mathcal{C}(f)$; nous devons prouver que g est un polynôme de f. Par hypothèse nous avons un vecteur cyclique que nous notons v. Nous avons un polynôme P (dépendant de g) tel que $g(v) = P(f)v$. Nous allons voir que $g = P(f)$. Soit $y \in E$ et un polynôme Q tel que $y = Q(f)v$; en notant que g commute avec $P(f)$ nous avons

$$g(y) = g\big(Q(f)v\big) = Q(f)\big(g(v)\big) = Q(f)\big(P(f)v\big) = P(f)Q(f)v = P(f)y. \tag{9.542}$$

Donc $g = P(f)$.

(3) implique (1) Nous avons les inégalités :

$$n \leqslant \dim\big(\mathcal{C}(f)\big) = \dim\big(\mathbb{K}[f]\big) = \deg(\mu) \leqslant \deg(\chi) = n. \tag{9.543}$$

La première est le lemme 9.232. Toutes les inégalités sont des égalités. En particulier $\deg(\mu) = n$, ce qui signifie que $\mu = \chi$ parce que μ est un polynôme diviseur de χ, de même degré que χ et unitaire tout comme χ.

□

Corollaire 9.234 ([83]).
En suivant les notations sur les extensions de corps de base de la section 9.13, l'endomorphisme $f\colon E \to F$ est cyclique si et seulement si l'endomorphisme $f_{\mathbb{L}}\colon E_{\mathbb{L}} \to F_{\mathbb{L}}$ est cyclique.

Démonstration. Nous savons par le théorème 9.233 qu'un endomorphisme est cyclique si et seulement si son polynôme minimal est égal à son polynôme caractéristique. Or par les propositions 9.214 et 9.215, nous savons que ces polynômes sont identiques pour f et pour $f_{\mathbb{L}}$. □

Théorème 9.235 (Similitude et extension de corps[83]).
Les applications linéaires $f, g\colon E \to E$ sont semblables si et seulement si $f_{\mathbb{L}}$ et $g_{\mathbb{L}}$ le sont.

Démonstration. En ce qui concerne le sens direct, s'il existe $m \in \mathrm{GL}(E)$ tel que $f = mgm^{-1}$ alors il suffit d'appliquer le lemme 9.210 pour avoir $f_{\mathbb{L}} = m_{\mathbb{L}} g_{\mathbb{L}} m_{\mathbb{L}}^{-1}$.

Nous considérons les invariants de similitude de f du théorème 9.221. Il existe une unique suite de polynômes unitaires μ_i $(i = 1, \dots, s)$ tels que $\mu_i \mid \mu_{i+1}$ et pour laquelle nous avons une décomposition $E = E_1 \oplus \dots \oplus E_s$ pour laquelle $f|_{E_i}\colon E_1 \to E_i$ est cyclique et de polynôme minimal μ.

Nous avons aussi $E_{\mathbb{L}} = (E_1)_{\mathbb{L}} \oplus \dots \oplus (E_s)_{\mathbb{L}}$ et les $(E_i)_{\mathbb{L}}$ sont stables sous $f_{\mathbb{L}}$ qui y sera également cyclique (corollaire 9.234). De plus le polynôme minimal de $f_{\mathbb{L}}|_{(E_i)_{\mathbb{L}}}$ est également μ_i.

Autrement dit, la suite μ_i est également la suite des invariants de similitude de $f_{\mathbb{L}}$. La remarque 9.222 nous permet de conclure que f et g sont semblables si et seulement si $f_{\mathbb{L}}$ et $g_{\mathbb{L}}$ le sont. □

9.16 Théorème de Burnside

Lemme 9.236.
Soit P, un polynôme sur $\mathbb{K}$. Une racine de P est une racine simple si et seulement si elle n'est pas racine de P'.

Théorème 9.237.
Toute représentation d'un groupe abélien d'exposant fini sur $\mathbb{C}^n$ a une image finie.

Démonstration. Dans le cas d'un groupe abélien, la démonstration est facile. Étant donné que G est d'exposant fini, il existe $\alpha \in \mathbb{N}^*$ tel que $g^\alpha = e$ pour tout $g \in G$. Le polynôme $P(X) = X^\alpha - 1$ est scindé à racines simples. En effet tout polynôme sur $\mathbb{C}$ est scindé. Le fait qu'il soit à racines simples provient du lemme 9.236 parce que si $a^\alpha = 1$, alors il n'est pas possible d'avoir $\alpha a^{\alpha-1} = 0$.

Par ailleurs $P(g) = 0$. Le fait que nous ayons un polynôme annulateur de g scindé à racines simples implique que g est diagonalisable (théorème 9.143). Le fait que G soit abélien montre qu'il existe une base de $\mathbb{C}^n$ dans laquelle tous les éléments de G sont diagonaux. Nous devons par conséquent montrer qu'il existe un nombre fini de matrices de la forme

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}. \tag{9.544}$$

Nous savons que $\lambda_i^\alpha = 1$ parce que $g^\alpha = \mathbb{1}$, par conséquent chacun des λ_i est une racine de l'unité dont il n'existe qu'un nombre fini. □

Théorème 9.238 (Burnside[34, 116])**.**
Un sous-groupe de $\mathrm{GL}(n, \mathbb{C})$ est fini si et seulement s'il est d'exposant[59] *fini.*

Démonstration. Soit G un sous-groupe de $\mathrm{GL}(n, \mathbb{C})$. Si G est fini, l'ordre de ses éléments divise $|G|$ (corollaire 3.55) au théorème de Lagrange et l'exposant est le PPCM qui est donc fini également.

Nous supposons maintenant que l'ordre de G est fini. Nous notons e l'exposant de G. En particulier, tous les éléments de G sont des racines du polynôme $X^e - 1$.

Générateurs Le groupe G est une partie de $\mathbb{M}(n, \mathbb{C})$ dont nous considérons l'algèbre engendrée (définition 4.35) $\mathcal{G}$. Soit $C_1, \dots, C_r$ une famille génératrice de $\mathcal{G}$ constituée d'éléments de G et la fonction

$$\begin{aligned} \tau\colon G &\to \mathbb{C}^r \\ A &\mapsto \big(\operatorname{Tr}(AC_1), \dots, \operatorname{Tr}(AC_r)\big). \end{aligned} \tag{9.545}$$

τ est injective Supposons que $\tau(A) = \tau(B)$. Alors pour tout générateur C_i nous avons $\operatorname{Tr}(AC_i) = \operatorname{Tr}(BC_i)$ et par linéarité de la trace, nous avons

$$\operatorname{Tr}(AM) = \operatorname{Tr}(BM) \tag{9.546}$$

pour tout $M \in G$. Notons par ailleurs

$$N = AB^{-1} - \mathbb{1}, \tag{9.547}$$

qui est diagonalisable parce que $AB^{-1} \in G$ et donc est annulé par le polynôme $X^e - 1$ qui est scindé à racines simples. Du coup AB^{-1} est diagonalisable ; posons $PAB^{-1}P^{-1} = D$, alors $P\big(AB^{-1} - \mathbb{1}\big)P^{-1} = D - \mathbb{1}$ qui est encore diagonale. Donc N est diagonalisable.

Par ailleurs nous avons

$$\begin{aligned} \operatorname{Tr}\big((AB^{-1})^p\big) &= \operatorname{Tr}\big(AB^{-1}(AB^{-1})^{p-1}\big) && && \text{(9.548a)} \\ &= \operatorname{Tr}\big(BB^{-1}(AB^{-1})^{p-1}\big) && (9.546) && \text{(9.548b)} \\ &= \operatorname{Tr}\big((AB^{-1})^{p-1}\big). && && \text{(9.548c)} \end{aligned}$$

59. Définition 3.15.

En continuant nous obtenons

$$\mathrm{Tr}\left((AB^{-1})^p\right) = \mathrm{Tr}(\mathbb{1}) = n. \tag{9.549}$$

D'autre part,

$$N^k = (AB^{-1} - \mathbb{1})^k = \sum_{p=0}^{k} \binom{p}{k} (-1)^{k-p} (AB^{-1})^p \tag{9.550}$$

En prenant la trace, et en tenant compte du fait que $\mathrm{Tr}\left((AB^{-1})^p\right) = n$,

$$\mathrm{Tr}(N^k) = \sum_{p=0}^{k} \binom{p}{k} (-1)^{k-p} n = n(1-1)^k = 0. \tag{9.551}$$

Donc la trace de N^k est nulle et le lemme 9.136 nous enseigne que N est alors nilpotente. Étant donné qu'elle est aussi diagonalisable, elle est nulle. Nous en concluons que $AB^{-1} = \mathbb{1}$ et donc que $A = B$. La fonction τ est donc injective.

Nombre fini de valeurs Les éléments de G sont annulés par $X^e - 1$ qui est un polynôme scindé à racines simples. Dons le polynôme minimal d'un élément de G est (a fortiori) scindé à racines simples et le théorème 9.143 nous assure alors que ces éléments sont diagonalisables. Du coup les valeur propres des matrices de G sont des racines $e$ièmes de l'unité. Par conséquent les traces des éléments de G ne peuvent prendre qu'un nombre fini de valeurs : toutes les sommes de n racines $e$ièmes de l'unité. Mais vu que les C_i sont dans G, nous avons

$$\mathrm{Image}(\tau) = \{\mathrm{Tr}(A) \text{ tel que } A \in G\}^r, \tag{9.552}$$

qui est un ensemble fini. Par conséquent G est fini parce que τ est injective.

□

9.16.1 Théorème de Lie-Kolchin

Contrairement à ce que l'on peut parfois croire, il n'est pas vrai que toute matrice à coefficient réel est diagonalisable, même pas sur $\mathbb{C}$. La raison est qu'une telle matrice peut très bien avoir des valeurs propres multiples.

Exemple 9.239
Le théorème 9.143 nous donne une façon simple de trouver des matrices non diagonalisables sur $\mathbb{C}$: il suffit que le polynôme minimal ne soit pas scindé à racines simples. Par exemple

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \tag{9.553}$$

dont le polynôme caractéristique est $\chi_A = (1-X)^2$. Ce polynôme n'a manifestement pas des racines simples. Nous pouvons faire le calcul explicite pour montrer que A n'est pas diagonalisable. D'abord l'unique valeur propre de A est 1 et nous pouvons sans peine résoudre

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \tag{9.554}$$

qui revient au système

$$\begin{cases} x + y = x & \text{(9.555a)} \\ y = y. & \text{(9.555b)} \end{cases}$$

La première équation donne directement $y = 0$. Le seul espace propre est de dimension 1 et est engendré par $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. △

La remarque 9.192 donne un exemple un peu plus avancé, qui montre la multiplicité algébrique et géométrique d'une racine d'un polynôme caractéristique.

Lemme 9.240 (Trigonalisation simultanée)**.**
Une famille de matrices de $\mathrm{GL}(n,\mathbb{C})$ *commutant deux à deux est simultanément trigonalisable.*

Démonstration. Commençons par enfoncer une porte ouverte par la proposition 9.150 : sur $\mathrm{GL}(n,\mathbb{C})$ toutes les matrices sont trigonalisables parce que tous les polynômes sont scindés.

Nous effectuons la démonstration par récurrence sur la dimension. Si $n = 1$ alors toues les matrice sont triangulaires et nous ne nous posons pas de questions. Nous supposons donc $n > 1$.

Soit la famille $(A_i)_{i\in I}$ dans $\mathrm{GL}(n,\mathbb{C})$ et A_0 un de ses éléments. Nous nommons $\lambda_1,\dots,\lambda_r$ les valeurs propres distinctes de A_0. Le théorème de décomposition primaire 9.193 nous donne la somme directe d'espaces caractéristiques [60]

$$E = F_{\lambda_1}(A_0)\oplus\dots\oplus F_{\lambda_r}(A_0). \tag{9.556}$$

Nous pouvons supposer que cette somme n'est pas réduite à un seul terme. En effet si tel était le cas, A_0 serait un multiple de l'identité parce que A_0 n'aurait qu'une seule valeur propre et les sommes dans la décomposition de Dunford 9.194(3) se réduisent à un seul terme (et $p_i = \mathrm{Id}$). En particulier les dimensions des espaces $F_\lambda(A_0)$ sont strictement plus petites que n.

Vu que tous les A_i commutent avec A_0, les espaces $F_\lambda(A_0)$ sont stables par les A_i et nous pouvons trigonaliser les A_i simultanément sur chacun des $F_\lambda(A_0)$ en utilisant l'hypothèse de récurrence. □

Théorème 9.241 (Lie-Kolchin[117])**.**
Tout sous-groupe connexe et résoluble de $\mathrm{GL}(n,\mathbb{C})$ *est conjugué à un groupe de matrices triangulaires.*

Démonstration. Soit G un sous-groupe connexe et résoluble de $\mathrm{GL}(n,\mathbb{C})$.

Si sous-espace non trivial stable par G Nous commençons par voir ce qu'il se passe s'il existe un sous-espace vectoriel non trivial V de $\mathbb{C}^n$ stabilisé par G. Pour cela nous considérons une base de $\mathbb{C}^n$ dont les premiers éléments forment une base de V (base incomplète, théorème 7.10). Les éléments de G s'écrivent, dans cette base,

$$\begin{pmatrix} g_1 & * \\ 0 & g_2 \end{pmatrix}. \tag{9.557}$$

Les matrices g_1 et g_2 sont carrés. Nous considérons alors l'application ψ définie par

$$\begin{aligned} \psi\colon G &\to \mathrm{GL}(V) \\ g &\mapsto g_1. \end{aligned} \tag{9.558}$$

Cela est un morphisme de groupes parce que

$$\begin{pmatrix} g_1 & * \\ 0 & g_2 \end{pmatrix}\begin{pmatrix} h_1 & * \\ 0 & h_2 \end{pmatrix} = \begin{pmatrix} g_1h_1 & * \\ 0 & g_2h_2 \end{pmatrix}, \tag{9.559}$$

de telle sorte que $\psi(gh) = \psi(g)\psi(h)$.

Le groupe $\psi(G)$ est connexe et résoluble. En effet $\psi(G)$ est connexe en tant qu'image d'un connexe par une application continue (proposition 6.86). Et il est résoluble en tant qu'image d'un groupe résoluble par un homomorphisme par la proposition 3.71. Vu que $\psi(G)$ est un sous-groupe résoluble et connexe de $\mathrm{GL}(V)$ et que la dimension de V est strictement plis petite que celle de $\mathbb{C}^n$, une récurrence sur la dimension indique que $\psi(G)$ est conjugué à un groupe de matrices triangulaires. C'est à dire qu'il existe une base de V dans laquelle toutes les matrices g_1 (avec $g\in G$) sont triangulaires supérieures.

On fait de même avec l'application $g\mapsto g_2$, ce qui donne une base du supplémentaire de V dans laquelle les matrices g_2 sont triangulaires.

En couplant ces deux bases, nous obtenons une base de $\mathbb{C}^n$ dans laquelle toutes les matrices (9.557) (c'est à dire toutes les matrices de G) sont triangulaires supérieures.

60. Définition 9.190.

Sinon Nous supposons à présent que $\mathbb{C}^n$ n'a pas de sous-espaces non triviaux stables sous G. Nous posons $m = \min\{k \text{ tel que } D^k(G) = \{e\}\}$, qui existe parce que G et résoluble et que sa suite dérivée termine sur e (proposition 3.70).

Si $m = 1$ Si $m = 1$ alors G est abélien et il existe une base de G dans laquelle toutes les matrices de G sont triangulaires (lemme 9.240). Le premier vecteur d'une telle base serait stable par G, mais comme nous avons supposé qu'il n'y avait pas de sous-espaces non triviaux stabilisés par G, il faut déduire que ce vecteur stable est à lui tout seul non trivial, c'est à dire que $n = 1$. Dans ce cas, le théorème est démontré.

Si $m > 1$ Nous devons maintenant traiter le cas où $m > 1$. Nous posons $H = D^{m-1}(G)$; cela est un sous-groupe normal et abélien de G. Encore une fois le résultat de trigonalisation simultanée 9.240 donne une base dans laquelle tous les éléments de H sont triangulaires. En particulier le premier élément de cette base est un vecteur propre commun à toutes les matrices de H.

Soit V le sous-espace engendré par tous les vecteurs propres communs de H. Nous venons de voir que V n'est pas vide. Nous allons montrer que V est stable par G. Soient $h \in H$, $v \in V$ et $g \in G$:

$$h\big(g(v)\big) = g\underbrace{g^{-1}hg}_{\in H}(v) = g(\lambda v) = \lambda g(v) \tag{9.560}$$

parce que v est vecteur propre de $g^{-1}hg$. Ce que le calcul (9.560) montre est que $g(v)$ est vecteur propre de h pour la valeur propre λ. Donc $g(v) \in V$ et V est stabilisé par G. Mais comme il n'existe pas d'espaces non triviaux stabilisés par G, nous en déduisons que $V = \mathbb{C}^n$. Donc tous les vecteurs de $\mathbb{C}^n$ sont vecteurs propres communs de H. Autrement dit on a une base de diagonalisation simultanée de H.

H est dans le centre de G Montrons à présent que H est dans le centre de G, c'est à dire que pour tout $g \in G$ et $h \in H$ il faut $ghg^{-1} = h$. D'abord ghg^{-1} est une matrice diagonale (parce que elle est dans H) ayant les mêmes valeurs propres que h. En effet si λ est valeur propre de ghg^{-1} pour le vecteur propre v, alors

$$(ghg^{-1})(v) = \lambda v \tag{9.561a}$$
$$h\big(g^{-1}v\big) = \lambda\big(g^{-1}v\big), \tag{9.561b}$$

c'est à dire que λ est également valeur propre de h, pour le vecteur propre $g^{-1}v$. Mais comme h a un nombre fini de valeurs propres, il n'y a qu'un nombre fini de matrices diagonales ayant les mêmes valeurs propres que h. L'ensemble $\mathbf{Ad}(G)h$ est donc un ensemble fini. D'autre part, l'application $g \mapsto g^{-1}hg$ est continue, et G est connexe, donc l'ensemble $\mathbf{Ad}(G)h$ est connexe. Un ensemble fini et connexe dans $\mathrm{GL}(n, \mathbb{C})$ est nécessairement réduit à un seul point. Cela prouve que $ghg^{-1} = h$ pour tout $g \in G$ et $h \in H$.

Espaces propres stables pour tout G Soit $h \in H$ et W un espace propre de h (ça existe non vide parce que H est triangularisé, voir plus haut). Alors nous allons prouver que W est stable pour tous les éléments de G. En effet si $w \in W$ avec $h(w) = \lambda w$ alors en permutant g et h,

$$hg(w) = g(hw) = \lambda g(w), \tag{9.562}$$

donc $g(w)$ est aussi vecteur propre de h pour la valeurs propre λ, c'est à dire que $g(w) \in W$. Vu que nous supposons que $\mathbb{C}^n$ n'a pas d'espaces invariants non triviaux, nous devons conclure que $W = \mathbb{C}^n$, c'est à dire que H est composé d'homothéties. C'est à dire que pour tout $h \in H$ nous avons $h = \lambda_h \mathbb{1}$.

Contradiction sur la minimalité de m Les éléments d'un groupe dérivé sont de déterminant 1 parce que $\det(g_1 g_2 g_1^{-1} g_2^{-1}) = 1$. Par conséquent pour tout h, le nombre λ_h est une racine n^{e}de l'unité. Vu qu'il n'y a qu'une quantité finie de racines n^{e}de l'unité, le groupe H est fini et connexe et donc une fois de plus réduit à un élément, c'est à dire $H = \{e\}$. Cela contredit la minimalité de m et donc produit une contradiction. Nous devons donc avoir $m = 1$.

Conclusion Nous avons vu que si $\mathbb{C}^n$ avait un sous-espace non trivial fixé par G alors le théorème était démontré. Par ailleurs si $\mathbb{C}^n$ n'a pas un tel sous-espace, soit $m = 1$ (et alors le théorème est également prouvé), soit $m > 1$ et alors on a une contradiction.

Bref, le théorème est prouvé sous peine de contradiction.

□

9.17 Formes bilinéaires et quadratiques

9.17.1 Généralités

Définition 9.242 ([118])**.**
Soit un espace vectoriel E et $\mathbb{F}$ un corps de caractéristique différente de 2. Une ***forme quadratique*** *sur E est une application $q\colon V \to \mathbb{F}$ pour laquelle il existe une forme bilinéaire symétrique $b\colon V \times V \to \mathbb{F}$ satisfaisant $q(x) = b(x, x)$ pour tout $x \in V$.*

L'ensemble des formes quadratiques réelles sur E est noté $Q(E)$.

Lemme 9.243.
Si q est une forme quadratique, il existe une unique forme bilinéaire b telle que $q(x) = b(x, x)$.

Démonstration. L'existence n'est pas en cause : c'est la définition d'une forme quadratique. Pour l'unicité, étant donné une forme quadratique, la forme bilinéaire b doit forcément vérifier l'**identités de polarisation** :

$$b(x, y) = \frac{1}{2}\big(q(x) + q(y) - q(x - y)\big). \tag{9.563}$$

Elle est donc déterminée par q. □

Notons la division par 2 qui est le pourquoi de la demande de la caractéristique différente de 2 pour $\mathbb{F}$ dans la définition de forme quadratique.

9.17.2 Topologie

La topologie considérée sur $Q(E)$ est celle de la norme

$$N(q) = \sup_{\|x\|_E=1} |q(x)|, \tag{9.564}$$

qui du point de vue de $S_n(\mathbb{R})$ est

$$N(A) = \sup_{\|x\|_E=1} |x^t A x|. \tag{9.565}$$

Notons que à droite, c'est la valeur absolue usuelle sur $\mathbb{R}$.

9.17.3 Matrice associée

Proposition 9.244.
Soit M, une matrice symétrique. Nous avons

(1) $\det M > 0$ *et* $\mathrm{Tr}(M) > 0$ *implique M définie positive,*

(2) $\det M > 0$ *et* $\mathrm{Tr}(M) < 0$ *implique M définie négative,*

(3) $\det M < 0$ *implique ni semi définie positive, ni définie négative*

(4) $\det M = 0$ *implique M semi-définie positive ou semi-définie négative.*

Si une base $\{e_i\}_{i=1,\dots,n}$ de l'espace vectoriel E est donnée, la **matrice associée** à la forme bilinéaire b sur E est la matrice d'éléments

$$B_{ij} = b(e_i, e_j). \tag{9.566}$$

Notons que la matrice associée à une forme bilinéaire (ou quadratique associée) est uniquement valable pour une base donnée. Si nous changeons de base, la matrice change. Cependant lorsque nous travaillons sur $\mathbb{R}^n$, la base canonique est tellement canonique que nous allons nous permettre de parler de «la» matrice associée à une forme bilinéaire.

Si B_{ij} est la matrice associée à la forme bilinéaire b alors la valeur de $b(u,v)$ se calcule avec la formule

$$b(x,y)=\sum_{i,j}B_{ij}x_iy_j \tag{9.567}$$

lorsque x_i et y_j sont les coordonnées de x et y dans la base choisie.

Proposition 9.245.
Soit $\{e_i\}$ une base de E. L'application

$$\begin{aligned}\phi\colon Q(E)&\to S(n,\mathbb{R})\\ q&\mapsto \big(b(e_i,e_j)\big)_{i,j}\end{aligned} \tag{9.568}$$

où b est forme bilinéaire associée à q est une bijection linéaire et continue.

Démonstration. Si $\phi(q)=\phi(q')$; alors

$$q(x)=\sum_{i,j}\phi(q)_{ij}x_ix_j=\sum_{i,j}\phi(q')_{ij}x_ix_j=q'(x). \tag{9.569}$$

Donc $q=q'$. L'application ϕ est donc injective

De plus elle est surjective parce que si $B\in S(n,\mathbb{R})$ alors la forme quadratique

$$q(x)=\sum_{i,j}B_{ij}x_ix_j \tag{9.570}$$

a évidemment B comme matrice associée. L'application ϕ est donc surjective.

Notre application ϕ est de plus linéaire parce que l'association d'une forme quadratique à la forme bilinéaire associée est linéaire.

En ce qui concerne la continuité, nous la prouvons en zéro en considérant une suite convergente $q_n\xrightarrow{Q(E)}0$. C'est à dire que

$$\sup_{\|x\|=1}|q_n(x)|\to 0. \tag{9.571}$$

Nous rappelons l'identité de polarisation :

$$b_n(x,y)=\frac{1}{2}\big(q_n(x-y)-q(x)-q(y)\big). \tag{9.572}$$

En ce qui concerne deux des trois termes, il n'y a pas de problèmes :

$$\big|\phi(q_n)_{ij}\big|=\big|b_n(e_i,e_j)\big|\leqslant\frac{1}{2}\big|b_n(e_i-e_j)\big|+\frac{1}{2}\big|q_n(e_i)\big|+\frac{1}{2}\big|q_n(e_j)\big|. \tag{9.573}$$

Si n est assez grand, nous avons tout de suite

$$\big|\phi(q_n)_{ij}\big|\leqslant\frac{1}{2}\big|q_n(e_i-e_j)\big|+\epsilon. \tag{9.574}$$

Nous définissons e_{ij} et α_{ij} de telle sorte que $e_i-e_j=\alpha_{ij}e_{ij}$ avec $\|e_{ij}\|=1$. Si $\alpha=\max\{\alpha_{ij},1\}$ alors nous avons

$$q_n(e_i-e_j)=\alpha_{ij}^2q_n(e_{ij})\leqslant\alpha^2q_n(e_{ij}). \tag{9.575}$$

Il suffit maintenant de prendre n assez grand pour avoir $\sup_{\|x\|=1}|q_n(x)|\leqslant\frac{\epsilon}{\alpha^2}$ pour avoir

$$\big|\phi(q_n)_{ij}\big|\leqslant\frac{\epsilon}{2}+\frac{\epsilon}{\alpha^2}. \tag{9.576}$$

□

Proposition 9.246.
Dans la base de diagonalisation de sa matrice associée, une forme quadratique a la forme

$$q(x) = \sum_i \lambda_i x_i^2 \tag{9.577}$$

où les λ_i sont les valeurs propres de la matrice associée à q.

Démonstration. Soit q une forme quadratique et b la forme bilinéaire associée. Si $\{f_i\}$ est une base de diagonalisation de la matrice de b alors dans cette base nous avons

$$q(x) = b(x,x) = \sum_{ij} x_i x_j b(f_i, f_j) = \sum_i \lambda_i x_i^2 \tag{9.578}$$

où les λ_i sont les valeurs propres de la matrice de b. □

Notons que si nous choisissons une autre base de diagonalisation, les λ_i ne changement pas (à part l'ordre éventuellement). Cela pour dire que nous nous permettrons de parler des **valeurs propres** d'une forme quadratique comme étant les valeurs propres de la matrice associée.

Proposition 9.247.
Une application linéaire est définie positive si et seulement si sa matrice associée l'est.

9.17.4 Dégénérescence

Soit b, une forme bilinéaire symétrique non dégénérée sur l'espace vectoriel E de dimension n sur $\mathbb{K}$ où $\mathbb{K}$ est un corps de caractéristique différente de 2. Nous notons q la forme quadratique associée.

Définition 9.248.
*Une forme bilinéaire est **non dégénérée** $b(x,z) = 0$ pour tout z implique $x = 0$.*

Lemme 9.249.
Soit b une forme bilinéaire non dégénérée. Si x et y sont tels que $b(x,z) = b(y,z)$ pour tout z, alors $x = y$.

Démonstration. C'est immédiat du fait de la linéarité en le premier argument et de la non-dégénérescence : si $b(x,z) - b(y,z) = 0$ alors

$$b(x - y, z) = 0 \tag{9.579}$$

pour tout z, ce qui implique $x - y = 0$. □

Proposition 9.250.
La forme bilinéaire b est non-dénénérée si et seulement si sa matrice associée est inversible.

Démonstration. Nous savons que la matrice associée est symétrique et qu'elle peut donc être diagonalisée (théorème 9.165). En nous plaçant dans une base de diagonalisation, nous devons prouver que la forme est non-dégénérée si et seulement si les éléments diagonaux de la matrice sont tous non nuls.

Écrivons $b(x,z)$ en choisissant pour z le vecteur de base e_k de composantes $(e_k)_j = \delta_{kj}$:

$$b(x, e_k) = \sum_{ij} x_i (e_k)_j = \sum_i b_{ik} x_i = b_{kk} x_k. \tag{9.580}$$

Si b est dégénérée et si x est un vecteur non nul (disons que la composante x_i est non nulle) de E tel que $b(x,z) = 0$ pour tout $z \in E$, alors $b_{ii} = 0$, ce qui montre que la matrice de b n'est pas inversible.

Réciproquement si la matrice de b est inversible, alors tous les b_{kk} sont différents de zéro, et le seul vecteur x tel que $b_{kk}x_k = 0$ pour tout k est le vecteur nul. □

Définition 9.251 (Isotropie)**.**
Un vecteur est ***isotrope*** *pour b s'il est perpendiculaire à lui-même ; en d'autres termes, x est isotrope si et seulement si $b(x,x)=0$. Un sous-espace $W\subset E$ est* ***totalement isotrope*** *si pour tout $x,y\in W$, nous avons $b(x,y)=0$.*

Le ***cône isotrope*** *de b est l'ensemble de ses vecteurs isotropes :*

$$C(b)=\{x\in E \text{ tel que } b(x,x)=0\}. \tag{9.581}$$

Nous introduisons quelque notations. D'abord pour $y\in E$ nous notons

$$\begin{aligned}\Phi_y\colon E&\to\mathbb{R}\\ x&\mapsto b(x,y)\end{aligned} \tag{9.582}$$

et ensuite

$$\begin{aligned}\Phi\colon E&\to E^*\\ y&\mapsto \Phi_y.\end{aligned} \tag{9.583}$$

Définition 9.252.
Le fait pour une forme bilinéaire b d'être dégénérée signifie que l'application Φ n'est pas injective. Le ***noyau*** *de la forme bilinéaire est celui de Φ, c'est à dire*

$$\ker(b)=\{z\in E \text{ tel que } b(z,y)=0\,\forall y\in E\}. \tag{9.584}$$

Autrement dit, $\ker(b)=E^\perp$ où le perpendiculaire est pris par rapport à b.

Notons tout de même que nous utilisons la notation $\perp$ même si b est dégénérée et éventuellement pas positive ; c'est à dire même si la formule $(x,y)\mapsto b(x,y)$ ne fournit pas un produit scalaire.

Proposition 9.253 ([119])**.**
Soit b une forme bilinéaire et symétrique. Alors

(1) $\ker(b)\subset C(b)$ *(cône d'isotropie, définition 9.251)*

(2) si b est positive alors $\ker(b)=C(b)$.

Démonstration. (1) Si $z\in\ker(b)$ alors pour tout $y\in E$ nous avons $b(z,y)=0$. En particulier pour $y=z$ nous avons $b(z,z,)=0$ et donc $z\in C(b)$.

(2) Soit b positive et $x\in C(b)$. Par l'inégalité de Cauchy-Schwarz (proposition 9.25) nous avons

$$|b(x,y)|\leqslant\sqrt{b(x,x)b(y,y)}=0. \tag{9.585}$$

Donc pour tout y nous avons $b(x,y)=0$.

□

9.17.5 Inégalité de Minkowski

Ce qui est couramment nommé «inégalité de Minkowski» est la proposition 24.26 dans les espaces L^p. Nous allons en donner ici un cas très particulier.

Proposition 9.254.
Si q est une forme quadratique sur $\mathbb{R}^n$ et si $x,y\in\mathbb{R}^n$ alors

$$\sqrt{q(x+y)}\leqslant\sqrt{q(x)}+\sqrt{q(y)}. \tag{9.586}$$

Démonstration. La proposition 9.246 nous permet de «diagonaliser» la forme quadratique q. Quitte à ne plus avoir une base orthonormale, nous pouvons renormaliser les vecteurs de base pour avoir

$$q(x)=\sum_i x_i^2. \tag{9.587}$$

Le résultat n'est donc rien d'autre que l'inégalité triangulaire pour la norme euclidienne usuelle, laquelle est démontrée dans la proposition 9.29. □

9.17.6 Ellipsoïde

Lemme 9.255.
Toute matrice peut être décomposée de façon unique en une partie symétrique et une partie antisymétrique. Cette décomposition est donnée par

$$S = \frac{M + M^t}{2}, \quad A = \frac{M - M^t}{2} \tag{9.588}$$

Démonstration. L'existence est une vérification immédiate de $S + A = M$ en utilisant (9.588). Pour l'unicité, si $S + A = S' + A'$ alors $S - S' = A - A'$. Mais $S - S'$ est symétrique et $A - A'$ est antisymétrique ; l'égalité implique l'annulation des deux membres, c'est à dire $S = S'$ et $A = A'$. □

Définition 9.256.
*Un **ellipsoïde** dans $\mathbb{R}^n$ centré en v est le lieu des points x vérifiant l'équation*

$$(x - v)^t M(x - v) = 1 \tag{9.589}$$

où M est une matrice symétrique strictement définie positive[61].

Lorsque nous parlons d'ellipsoïde plein, *il suffit de changer l'égalité en une inégalité.*

Une autre façon d'écrire la relation (9.589) est d'écrire $\langle (x-v), M(x,v)\rangle$ en utilisant le produit scalaire.

Remarque 9.257.
Le fait que M soit symétrique n'est pas tout à fait obligatoire ; la chose important est que toutes les valeurs propres soient strictement positives. En effet si M a toutes ses valeurs propres strictement positives, nous nommons S la partie symétrique de M et A la partie antisymétrique (lemme 9.255). Alors pour tout $x \in \mathbb{R}^n$ nous avons

$$x^t Ax = \langle x, Ax\rangle = \langle A^t x, x\rangle = -\langle Ax, x\rangle = -\langle x, Ax\rangle, \tag{9.590}$$

donc $x^t Ax = 0$. L'équation $x^t Mx = 1$ est donc équivalente à $x^t Sx = 1$ (elles ont les mêmes solutions).

De plus S reste strictement définie positive parce que pour tout $x \in \mathbb{R}^n$ nous avons

$$0 < x^t Mx = x^t Sx. \tag{9.591}$$

Proposition 9.258.
Si $\mathcal{E}$ est un ellipsoïde centrée à l'origine, il existe une base de $\mathbb{R}^n$ dans laquelle son équation est :

$$\sum_{i=1}^{n} \frac{x_i^2}{a_i^2} = 1. \tag{9.592}$$

Démonstration. Nous avons une matrice symétrique strictement définie positive S telle que l'équation soit $\langle x, Sx\rangle = 1$. Le théorème spectral 9.165 nous fournit une base orthonormale $\{e_i\}$ dans laquelle $Se_i = \lambda_i e_i$ avec $\lambda_i > 0$. En substituant dans l'équation $\langle x, Sx\rangle = 1$ nous trouvons l'équation

$$\sum_i \lambda_i x_i^2 = 1. \tag{9.593}$$

En posant $a_i = \frac{1}{\sqrt{\lambda_i}}$, nous trouvons le résultat. Cette définition des a_i est toujours possible parce que $\lambda_i > 0$. □

Corollaire 9.259.
Un ellipsoïde plein centré en l'origine admet une équation de la forme $q(x) \leqslant 1$ où q est une forme quadratique strictement définie positive.

61. Définition 9.167.

Pour rappel de notation, l'ensemble des formes quadratiques strictement définies positives sur l'espace vectoriel E est noté $Q^{++}(E)$.

Démonstration. Soit $\{e_i\}$ une base de $\mathbb{R}^n$ telle que l'ellipsoïde $\mathcal{E}$ ait pour équation

$$\sum_{i=1}^{n} \frac{x_i^2}{a_i^2} \leqslant 1. \tag{9.594}$$

Nous considérons la forme quadratique

$$\begin{aligned} q\colon \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \sum_{i=1}^{n} \frac{\langle x, e_i\rangle^2}{a_i^2}. \end{aligned} \tag{9.595}$$

Nous avons évidemment $\mathcal{E} = \{x \in \mathbb{R}^n \text{ tel que } q(x) \leqslant 1\}$. De plus la forme q est strictement définie positive parce que dès que $x \neq 0$, au moins un des produits scalaires $\langle x, e_i\rangle$ est non nul et $q(x) > 0$. □

9.18 Théorème spectral auto-adjoint

Définition 9.260.
*Si E est un espace euclidien, un endomorphisme $f\colon E \to E$ est **auto-adjoint** si pour tout $x, y \in E$ nous avons $\langle x, f(y)\rangle = \langle f(x), y\rangle$.*

L'ensemble des opérateurs auto-adjoints de E est noté $\mathrm{S}(E)$. Cette notation provient du fait que dans $\mathbb{R}^n$ muni du produit scalaire usuel, les opérateurs auto-adjoints sont les matrices symétriques.

Théorème 9.261 (Théorème spectral auto-adjoint)**.**
Un endomorphisme auto-adjoint d'un espace euclidien

(1) est diagonalisable dans une base orthonormée,

(2) a son spectre réel.

Démonstration. Nous procédons par récurrence sur la dimension de E, et nous commençons par $n = 1$ [62]. Soit donc $f\colon E \to E$ avec $\langle f(x), y\rangle = \langle x, f(y)\rangle$. Étant donné que f est également linéaire, il existe $\lambda \in \mathbb{R}$ tel que $f(x) = \lambda x$ pour tout $x \in E$. Tous les vecteurs de E sont donc vecteurs propres de f.

Passons à la récurrence. Nous considérons $\dim(E) = n + 1$ et $f \in \mathrm{S}(E)$. Nous considérons la forme bilinéaire symétrique Φ_f et la forme quadratique associée ϕ_f. Pour rappel,

$$\Phi_f(x, y) = \langle x, f(y)\rangle \tag{9.596a}$$

$$\phi_f(x) = \Phi_f(x, x). \tag{9.596b}$$

Et nous allons laisser tomber les indices f pour noter simplement Φ et ϕ. Étant donné que $\overline{B(0,1)}$ est compacte et que ϕ est continue, il existe $x_0 \in \overline{B(0,1)}$ tel que

$$\lambda = \phi(x_0) = \sup_{x \in \overline{B(0,1)}} \phi(x). \tag{9.597}$$

Notons aussi que $\|x_0\| = 1$: le maximum est pris sur le bord. Nous posons

$$g = \lambda \mathrm{Id} - f \tag{9.598}$$

ainsi que

$$\Phi_1(x, y) = \langle x, g(y)\rangle. \tag{9.599}$$

62. Dans [2], l'auteur commence avec $n = 0$ mais moi je n'en ai pas le courage..

Cela est une forme bilinéaire et symétrique parce que

$$\Phi_1(y,x) = \langle y, g(x)\rangle = \langle g(y), x\rangle = \langle x, g(y)\rangle = \Phi_1(x,y) \tag{9.600}$$

où nous avons utilisé le fait que g était auto-adjoint et la symétrie du produit scalaire. De plus Φ_1 est semi-définie positive parce que

$$\Phi_1(x,x) = \langle x, \lambda x - f(x)\rangle = \lambda\|x\|^2 - \phi(x). \tag{9.601}$$

Vu que λ est le maximum, nous avons tout de suite $\Phi_1(x) \geqslant 0$ tant que $\|x\| = 1$. Et si x n'est pas de norme 1, c'est le même prix parce qu'on se ramène à $\|x\| = 1$ en multipliant par un nombre positif. Attention cependant :

$$\Phi_1(x_0,x_0) = \lambda\|x_0\|^2 - \phi(x_0) = 0. \tag{9.602}$$

Donc Φ_1 a un noyau contenant x_0 par la proposition 9.253. Nous en déduisons que $\text{Image}(g) \neq E$ en effet, $x_0 \in \text{Image}(g)^\perp$, mais nous avons la proposition 7.41 sur les dimensions :

$$\dim E = \dim(\text{Image}(g)) + \dim(\text{Image}(g)^\perp). \tag{9.603}$$

Vu que $\text{Image}(g)^\perp$ est un espace vectoriel non réduit à $\{0\}$, la dimension de $\text{Image}(g)$ ne peut pas être celle de E. L'endomorphisme g n'étant pas surjectif, il ne peut pas être injectif non plus parce que nous sommes en dimension finie ; il existe donc $e_1 \in E$ tel que $g(e_1) = 0$ et tant qu'à faire nous choisissons $\|e_1\| = 1$ (ici la norme est bien celle de l'espace euclidien considéré). Par définition,

$$f(e_1) = \lambda e_1, \tag{9.604}$$

c'est à dire que $\lambda \in \text{Spec}(f)$. Et ϕ étant une forme quadratique réelle nous avons $\lambda \in \mathbb{R}$.

Nous posons à présent $H = \text{Span}\{e_1\}^\perp$. C'est un sous-espace stable par f parce que si $x \in H$ alors

$$\langle e_1, f(x)\rangle = \langle f(e_1 j), x\rangle = \lambda\langle e_1, x\rangle = 0. \tag{9.605}$$

Nous pouvons donc considérer la restriction de f à H : $f_H\colon H \to H$. Cet endomorphisme est bilinéaire et symétrique sur l'espace H de dimension inférieure à celle de E, donc la récurrence nous donne une base orthonormée

$$\{e_2, \dots, e_n\} \tag{9.606}$$

de vecteurs propres de f_H. De plus les valeurs propres sont réelles, toujours par récurrence. Donc

$$\text{Spec}(f) = \{\lambda\} \cup \text{Spec}(f_H) \subset \mathbb{R}. \tag{9.607}$$

Notons pour être complet que si $i \geqslant 2$ alors

$$\langle e_1, e_i\rangle = 0 \tag{9.608}$$

parce que le vecteur e_i est par construction choisit dans l'espace $H = e_1^\perp$. Nous avons donc bien une base orthonormée de E construite sur des vecteurs propres de f. □

Corollaire 9.262.
Soit E un espace vectoriel ainsi que ϕ et ψ des formes quadratiques sur E avec ψ définie positive. Alors il existe une base ψ-orthonormale dans laquelle ϕ est diagonale.

Démonstration. Il suffit de considérer l'espace euclidien E muni du produit scalaire $\langle x, y\rangle = \psi(x,y)$. Ensuite nous diagonalisons la matrice (symétrique) de ϕ pour ce produit scalaire à l'aide du théorème 9.261. □

Définition 9.263.
*Dans le cas de $V = \mathbb{R}^m$ nous avons un produit scalaire canonique. Soient u et v, deux vecteurs de $\mathbb{R}^m$. Le **produit scalaire** de u et v, noté $\langle u, v\rangle$ ou $u \cdot v$ est le réel*

$$\langle u, v\rangle = \sum_{k=1}^{m} u_k v_k = u_1 v_1 + u_2 v_2 + \cdots + u_m v_n. \tag{9.609}$$

Calculons par exemple le produit scalaire de deux vecteurs de la base canonique : $\langle e_i, e_j \rangle$. En utilisant la formule de définition et le fait que $(e_i)_k = \delta_{ik}$, nous avons

$$\langle e_i, e_j \rangle = \sum_{k=1}^{m} \delta_{ik}\delta_{jk}. \tag{9.610}$$

Nous pouvons effectuer la somme sur k en remarquant qu'à cause du δ_{ik}, seul le terme avec $k = i$ n'est pas nul. Effectuer la somme revient donc à remplacer tous les k par des i :

$$\langle e_i, e_j \rangle = \delta_{ii}\delta_{ji} = \delta_{ji}. \tag{9.611}$$

Une des propriétés intéressantes du produit scalaire est qu'il permet de décomposer un vecteur dans une base, comme nous le montre la proposition suivante.

Proposition 9.264.
Si nous notons v_i les composantes du vecteur v, c'est à dire si $v = \sum_{i=1}^{m} v_i e_i$, alors nous avons $v_j = \langle v, e_j \rangle$.

Démonstration.

$$v \cdot e_j = \sum_{i=1}^{m} \langle v_i e_i, e_j \rangle = \sum_{i=1}^{m} v_i \langle e_i, e_j \rangle = \sum_{i=1}^{m} v_i \delta_{ij} \tag{9.612}$$

En effectuant la somme sur i dans le membre de droite de l'équation (9.612), tous les termes sont nuls sauf celui où $i = j$; il reste donc

$$v \cdot e_j = v_j. \tag{9.613}$$

□

Le produit scalaire ne dépend en réalité pas de la base orthogonale choisie.

Lemme 9.265.
Si $\{e_i\}$ est la base canonique, et si $\{f_i\}$ est une autre base orthonormale, alors si u et v sont deux vecteurs de $\mathbb{R}^m$, nous avons

$$\sum_i u_i v_j = \sum_i u'_i v'_j \tag{9.614}$$

où u_i sont les composantes de u dans la base $\{e_i\}$ et u'_i sont celles dans la base $\{f_i\}$.

Démonstration. La preuve demande un peu d'algèbre linéaire. Étant donné que $\{f_i\}$ est une base orthonormale, il existe une matrice A orthogonale ($AA^t = \mathbb{1}$) telle que $u'_i = \sum_j A_{ij} u_j$ et idem pour v. Nous avons alors

$$\begin{aligned}
\sum_i u'_i v'_j &= \sum_i \left(\sum_j A_{ij} u_j \right) \left(\sum_k A_{ik} v_k \right) \\
&= \sum_{ijk} A_{ij} A_{ik} u_j v_k \\
&= \sum_{jk} \underbrace{\sum_i (A^t)_{ji} A_{ik}}_{=\delta_{jk}} u_j v_k \\
&= \sum_{jk} \delta_{jk} u_j v_k \\
&= \sum_k u_j v_k.
\end{aligned} \tag{9.615}$$

□

Cette proposition nous permet de réellement parler du produit scalaire entre deux vecteurs de façon intrinsèque sans nous soucier de la base dans laquelle nous regardons les vecteurs.

Nous dirons que deux vecteurs sont **orthogonaux** lorsque leur produit scalaire est nul. Nous écrivons que $u \perp v$ lorsque $\langle u, v \rangle = 0$.

Définition 9.266.
*La **norme euclidienne** d'un élément de $\mathbb{R}^m$ est définie par $\|u\| = \sqrt{u \cdot u}$.*

Cette définition est motivée par le fait que le produit scalaire $u \cdot u$ donne exactement la norme usuelle donnée par le théorème de Pythagore :

$$u \cdot u = \sum_{i=1}^{m} u_i u_i = \sum_{i=1}^{m} u_i^2 = u_1^2 + u_2^2 + \cdots + u_m^2. \tag{9.616}$$

Le fait que $e_i \cdot e_j = \delta_{ij}$ signifie que la base canonique est **orthonormée**, c'est à dire que les vecteurs de la base canonique sont orthogonaux deux à deux et qu'ils ont tout 1 comme norme.

Lemme 9.267.
Pour tout $u \in \mathbb{R}^m$, il existe un $\xi \in \mathbb{R}^m$ tel que $\|u\| = \xi \cdot u$ et $\|\xi\| = 1$.

Démonstration. Vérifions que le vecteur $\xi = u/\|u\|$ ait les propriétés requises. D'abord $\|\xi\| = 1$ parce que $u \cdot u = \|u\|^2$. Ensuite

$$\xi \cdot u = \frac{u \cdot u}{\|u\|} = \frac{\|u\|^2}{\|u\|} = \|u\|. \tag{9.617}$$

□

9.19 Mini introduction au produit tensoriel

9.19.1 Définitions

Soit E, un espace vectoriel de dimension finie. Si α et β sont deux formes linéaires sur un espace vectoriel E, nous définissons $\alpha \otimes \beta$ comme étant la 2-forme donnée par

$$(\alpha \otimes \beta)(u, v) = \alpha(u)\beta(v). \tag{9.618}$$

Si a et b sont des vecteurs de E, ils sont vus comme des formes sur E via le produit scalaire et nous avons

$$(a \otimes b)(u, v) = (a \cdot u)(b \cdot v). \tag{9.619}$$

Cette dernière équation nous incite à pousser un peu plus loin la définition de $a \otimes b$ et de simplement voir cela comme la matrice de composantes

$$(a \otimes b)_{ij} = a_i b_j. \tag{9.620}$$

Cette façon d'écrire a l'avantage de ne pas demander de se souvenir qui est une vecteur ligne, qui est un vecteur colonne et où il faut mettre la transposée. Évidemment $(a \otimes b)$ est soit ab^t soit $a^t b$ suivant que a et b soient ligne ou colonne.

9.19.2 Application d'opérateurs

Lemme 9.268.
Soient $x, y \in E$ et A, B deux opérateurs linéaires sur E vus comme matrices. Alors

$$(Ax \otimes By) = A(x \otimes y)B^t. \tag{9.621}$$

Démonstration. Calculons la composante ij de la matrice $(Ax \otimes By)$. Nous avons

$$\begin{aligned}
(Ax \otimes By)_{ij} &= (Ax)_i (By)_j && \text{(9.622a)}\\
&= \sum_{kl} A_{ik} x_k B_{jl} y_l && \text{(9.622b)}\\
&= A_{ik} (x \otimes y)_{kl} B_{jl} && \text{(9.622c)}\\
&= \big(A(x \otimes y)B^t\big)_{ij}. && \text{(9.622d)}
\end{aligned}$$

□

9.20 Méthode de Gauss pour résoudre des systèmes d'équations linéaires

Pour résoudre un système d'équations linéaires, on procède comme suit :

(1) Écrire le système sous forme matricielle.

$$\text{p.ex. } \begin{cases} 2x+3y &= 5 \\ x+2y &= 4 \end{cases} \Leftrightarrow \left(\begin{array}{cc|c} 2 & 3 & 5 \\ 1 & 2 & 4 \end{array}\right)$$

(2) Se ramener à une matrice avec un maximum de 0 dans la partie de gauche en utilisant les transformations admissibles :

(a) Remplacer une ligne par elle-même + un multiple d'une autre ;

$$\text{p.ex. } \left(\begin{array}{cc|c} 2 & 3 & 5 \\ 1 & 2 & 4 \end{array}\right) \overset{L_1-2.L_2 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 4 \end{array}\right)$$

(b) Remplacer une ligne par un multiple d'elle-même ;

$$\text{p.ex. } \left(\begin{array}{cc|c} 0 & -1 & -3 \\ 1 & 2 & 4 \end{array}\right) \overset{-L_1 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 2 & 4 \end{array}\right)$$

(c) Permuter des lignes.

$$\text{p.ex. } \left(\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 0 & -2 \end{array}\right) \overset{L_1 \mapsto L_2' \text{ et } L_2 \mapsto L_1'}{\Longrightarrow} \left(\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 3 \end{array}\right)$$

(3) Retransformer la matrice obtenue en système d'équations.

$$\text{p.ex. } \left(\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 3 \end{array}\right) \Leftrightarrow \begin{cases} x &= -2 \\ y &= 3 \end{cases}$$

Remarque 9.269. — Si on obtient une ligne de zéros, on peut l'enlever :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \Leftrightarrow \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \end{array}\right)$$

— Si on obtient une ligne de zéros suivie d'un nombre non-nul, le système d'équations n'a pas de solution :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 3 & 4 & -2 & 2 \\ 4 & -1 & 3 & 0 \\ 0 & 0 & 0 & 7 \end{array}\right) \Leftrightarrow \begin{cases} \cdots \\ \cdots \\ 0x+0y+0z = 7 \end{cases} \Rightarrow \textbf{Impossible}$$

— Si on moins d'équations que d'inconnues, alors il y a une infinité de solutions qui dépendent d'un ou plusieurs paramètres :

$$\text{p.ex. } \left(\begin{array}{ccc|c} 1 & 0 & -2 & 2 \\ 0 & 1 & 3 & 0 \end{array}\right) \Leftrightarrow \begin{cases} x-2z = 2 \\ y+3z = 0 \end{cases} \Leftrightarrow \begin{cases} x = 2+2\lambda \\ y = -3\lambda \\ z = \lambda \end{cases}$$

Chapitre 10

Espaces vectoriels normés

10.1 Équivalence des normes

Au premier coup d'œil, les notions dont nous parlons dans ce chapitre ont l'air très générales. Nous prenons en effet n'importe quel espace vectoriel V de dimension finie, et nous le munissons de n'importe quelle norme (rien que dans $\mathbb{R}^m$ nous en avons défini une infinité par l'équation (6.21)). À partir de ces données, nous définissons les boules, la topologie, l'adhérence, etc.

10.1.1 En dimension finie

Dans $\mathbb{R}^n$, les normes $\|.\|_{L^1}$, $\|.\|_{L^2}$ et $\|.\|_\infty$ ne sont pas égales. Cependant elles ne sont pas complètement indépendante au sens où l'on sent bien que si un vecteur sera grand pour une norme, il sera également grand pour les autres normes ; les normes «vont dans le même sens». Cette notion est précisée par le concept de norme équivalente.

Définition 10.1.
*Deux normes N_1 et N_2 sur $\mathbb{R}^m$ sont **équivalentes** s'il existe deux nombres réels strictement positifs k_1 et k_2 tels que*

$$k_1 N_1(x) \leqslant N_2(x) \leqslant k_2 N_1(x), \tag{10.1}$$

pour tout x dans $\mathbb{R}^m$. Dans ce cas nous écrivons que $N_1 \sim N_2$.

Il est possible de démontrer que cette notion est une relation d'équivalence (définition 2.15) sur l'ensemble des normes existantes sur $\mathbb{R}^m$.

Proposition 10.2.
Pour $\mathbb{R}^N$, nous avons les équivalences de normes $\|.\|_{L^1} \sim \|.\|_{L^2}$, $\|.\|_{L^1} \sim \|.\|_\infty$ et $\|.\|_{L^2} \sim \|.\|_\infty$. Plus précisément nous avons les inégalités

(1) $\|x\|_2 \leqslant \|x\|_1 \leqslant \sqrt{n}\|x\|_2$

(2) $\|x\|_\infty \leqslant \|x\|_1 \leqslant n\|x\|_\infty$

(3) $\|x\|_\infty \leqslant \|x\|_2 \leqslant \sqrt{n}\|x\|_\infty$

Démonstration. En mettant au carré la première inégalité nous voyons que nous devons vérifier l'inégalité

$$|x_1|^2 + \cdots + |x_n|^2 \leqslant \big(|x_1| + \cdots + |x_n|\big)^2 \tag{10.2}$$

qui est vraie parce que le membre de droite est égal au carré de chaque terme plus les double produits. La seconde inégalité provient de l'inégalité de Cauchy-Schwarz (théorème 9.25) sur les vecteurs

$$v = \begin{pmatrix} 1/n \\ \vdots \\ 1/n \end{pmatrix}, \quad w = \begin{pmatrix} |x_1| \\ \vdots \\ |x_n| \end{pmatrix}. \tag{10.3}$$

Nous trouvons

$$\frac{1}{n}\sum_i |x_i| \leqslant \sqrt{b \cdot \frac{1}{n}}\sqrt{\sum_i |x_i|^2}, \tag{10.4}$$

et par conséquent

$$\sum_i |x_i| \leqslant \sqrt{n}\|x\|_2. \tag{10.5}$$

La première inégalité de (3) se démontre en remarquant que si a et b sont positifs, $a \leqslant \sqrt{a^2+b}$. En appliquant cela à $a = \max_i |x_i|$, nous avons

$$\max_i |x_i| \leqslant \sqrt{|x_1|^2 + \cdots + |x_n|^2} \tag{10.6}$$

parce que $\max_i |x_i|$ est évidemment un des termes de la somme. Pour la seconde inégalité de (3), nous avons

$$\sqrt{\sum_k |x_k|^2} \leqslant \left(\sum_k \max_i |x_i|^2\right)^{1/2} = \sqrt{n}\|x\|_\infty. \tag{10.7}$$

Pour obtenir cette inégalité, nous avons remplacé tous les termes $|x_k|$ par le maximum. □

En réalité, toutes les normes $\|.\|_{L^p}$ et $\|.\|_\infty$ sont équivalentes et, plus généralement, nous avons le résultat suivant, très étonnant à première vue, et en réalité assez difficile à prouver :

Théorème 10.3 ([120])**.**
Sur un espace vectoriel de dimension finie, toutes les normes (pas seulement les normes L^p que nous avons définies sur $\mathbb{R}^m$) sont équivalentes.

Corollaire 10.4.
Soit V un espace vectoriel de dimension finie et $\|.\|_1$, $\|.\|_2$ deux normes sur V. Alors l'identité $\mathrm{Id}\colon V \to V$ *est un isomorphisme d'espace topologique* $(V, \|.\|_1) \to (V, \|.\|_2)$.

De plus les ouverts sont les mêmes : une partie de V est ouverte dans $(V, \|.\|_1)$ si et seulement si elle est ouverte dans $(V, \|.\|_2)$.

Le théorème d'équivalence de norme sera utilisé pour montrer que l'ensemble des formes quadratiques non dégénérées de signature (p, q) est ouvert dans l'ensemble des formes quadratiques, proposition 18.101. Plus généralement il est utilisé à chaque fois que l'on fait de la topologie sur les espaces de matrices en identifiant $\mathbb{M}(n, \mathbb{R})$ à $\mathbb{R}^{n^2}$, pour se rassurer en se disant que ce qu'on fait ne dépend pas de la norme choisie.

10.1.2 Contre-exemple en dimension infinie

Lorsque nous considérons des espaces vectoriels de dimension infinie, les choses ne sons plus aussi simples. Nous voyons ici sur l'exemple de l'espace des polynômes que le théorème 10.3 n'est plus valable si on enlève l'hypothèse de dimension finie.

On considère l'ensemble des fonctions polynômiales à coefficients réels sur l'intervalle $[0, 1]$.

$$\mathcal{P}_{\mathbb{R}}([0,1]) = \{p : [0,1] \to \mathbb{R} \,|\, p : x \mapsto a_0 + a_1 x + a_2 x^2 + \ldots,\ a_i \in \mathbb{R},\ \forall i \in \mathbb{N}\}. \tag{10.8}$$

Cet ensemble, muni des opérations usuelles de somme entre polynômes et multiplications par les scalaires, est un espace vectoriel.

Sur $\mathcal{P}(\mathbb{R})$ on définit les normes suivantes

$$\begin{aligned} \|p\|_\infty &= \sup_{x\in[0,1]} \{p(x)\}, \\ \|p\|_1 &= \int_0^1 |p(x)|\, dx, \\ \|p\|_2 &= \left(\int_0^1 |p(x)|^2\, dx\right)^{1/2}. \end{aligned} \tag{10.9}$$

Les inégalités suivantes sont immédiates

$$\begin{aligned}\|p\|_1 &= \int_0^1 |p(x)|\,dx \leqslant \|p\|_\infty,\\ \|p\|_2 &= \left(\int_0^1 |p(x)|^2\,dx\right)^{1/2} \leqslant \|p\|_\infty,\end{aligned} \tag{10.10}$$

mais la norme $\|\cdot\|_\infty$ n'est équivalente ni à $\|\cdot\|_1$, ni à $\|\cdot\|_2$. Soit $p_k(x) = x^k$. Alors

$$\begin{aligned}\|p_k\|_\infty &= 1,\\ \|p_k\|_1 &= \int_0^1 x^k\,dx = \frac{1}{k+1},\\ \|p_k\|_2 &= \left(\int_0^1 x^{2k}\,dx\right)^{1/2} = \sqrt{\frac{1}{2k+1}}.\end{aligned} \tag{10.11}$$

Pour $k \to \infty$ les normes $\|p_k\|_1$, $\|p_k\|_2$ tendent vers zéro, alors que la norme $\|p_k\|_\infty$ est constante, donc les normes ne sont pas équivalentes parce que il n'existe pas un nombre positif m tel que

$$\begin{aligned}m\|p_k\|_\infty &\leqslant \|p_k\|_1,\\ m\|p_k\|_\infty &\leqslant \|p_k\|_2,\end{aligned} \tag{10.12}$$

uniformément pour tout k dans $\mathbb{N}$.

10.2 Norme opérateur

La proposition suivante donne une norme (au sens de la définition 6.31) sur $\mathcal{L}(V,W)$ dès que V et W sont des espaces vectoriels normés.

Proposition-définition 10.5 (Norme opérateur[121], thème 8)**.**
Soit une application linéaire $T\colon V \to W$, et le nombre

$$\|T\|_{\mathcal{L}} = \sup_{\substack{x\in V\\ x\neq 0}} \frac{\|T(x)\|_W}{\|x\|_V}. \tag{10.13}$$

(1) Si V est de dimension finie, alors $\|T\|_{\mathcal{L}} < \infty$.

(2) L'application $T \mapsto \|T\|_{\mathcal{L}}$ est une norme sur l'espace vectoriel des applications linéaires $V \to W$.

(3) Nous avons la formule

$$\|T\|_{\mathcal{L}} = \sup_{x\in V} \frac{\|T(x)\|_W}{\|x\|_V} = \sup_{\|x\|_V=1} \|T(x)\|_W \tag{10.14}$$

*Le nombre $\|T\|_{\mathcal{L}}$ est la **norme opérateur** de T. Nous disons que cette norme est **subordonnée** aux normes choisies sur V et W.*

Démonstration. Si V est de dimension finie alors l'ensemble $\{\|x\| = 1\}$ est compact par le théorème de Borel-Lebesgue 6.123. Alors la fonction

$$x \mapsto \frac{\|T(x)\|}{\|x\|} \tag{10.15}$$

est une fonction continue sur un compact. Le corollaire 6.130 nous dit alors qu'elle est bornée. Le supremum est donc un nombre réel fini.

Nous vérifions que l'application $\|.\|$ de $\mathcal{L}(V,W)$ dans $\mathbb{R}$ ainsi définie est effectivement une norme.

(1) $\|T\|_{\mathcal{L}} = 0$ signifie que $\|T(x)\| = 0$ pour tout x dans V. Comme $\| \cdot \|_W$ est une norme nous concluons que $T(x) = 0_n$ pour tout x dans V, donc T est l'application nulle.

(2) Pour tout a dans $\mathbb{R}$ et tout T dans $\mathcal{L}(V, W)$ nous avons

$$\|aT\|_{\mathcal{L}} = \sup_{\|x\|_V \leqslant 1} \|aT(x)\|_W = |a| \sup_{\|x\|_V \leqslant 1} \|T(x)\|_W = |a|\|T\|_{\mathcal{L}}. \tag{10.16}$$

(3) Pour tous T_1 et T_2 dans $\mathcal{L}(V, W)$ nous avons

$$\begin{aligned}\|T_1 + T_2\|_{\mathcal{L}} &= \sup_{\|x\| \leqslant 1} \|T_1(x) + T_2(x)\| \leqslant \\ &\leqslant \sup_{\|x\| \leqslant 1} \|T_1(x)\| + \sup_{\|x\| \leqslant 1} \|T_2(x)\| \\ &= \|T_1\|\|T_2\|.\end{aligned}$$

Enfin nous prouvons la formule alternative (10.14). Nous allons montrer que les ensembles sur lesquels ont prend le supremum sont en réalité les mêmes :

$$\underbrace{\left\{ \frac{\|Ax\|}{\|x\|} \right\}_{x \neq 0}}_{A} = \underbrace{\{\|Ax\| \text{ tel que } \|x\| = 1\}}_{B}. \tag{10.17}$$

Attention : ce sont des sous-ensembles de réels ; pas de sous-ensembles de $\mathbb{M}(\mathbb{R})$ ou des sous-ensembles de $\mathbb{R}^n$.

Pour la première inclusion, prenons un élément de A, et prouvons qu'il est dans B. C'est à dire que nous prenons $x \in V$ et nous considérons le nombre $\|Ax\|/\|x\|$. Le vecteur $y = x/\|x\|$ est un vecteur de norme 1, donc la norme de Ay est un élément de B, mais

$$\|Ay\| = \frac{\|Ax\|}{\|x\|}. \tag{10.18}$$

Nous avons donc $A \subset B$.

L'inclusion $B \subset A$ est immédiate. □

En d'autres termes, il y a autant de normes opérateur sur $\mathcal{L}(E, F)$ qu'il y a de paires de choix de normes sur E et F. En particulier, cela donne lieu à toutes les normes $\|A\|_p$ qui correspondent aux normes $\|.\|_p$ sur $\mathbb{R}^n$.

Exemple 10.6

Voyons la norme opérateur subordonnée à la norme $\|x\|_\infty = \max_i |x_i|$ sur $\mathbb{C}^n$. Par définition (et surtout par la propriété 10.5(3)),

$$\|A\|_\infty = \sup_{\|x\|_\infty = 1} = \|Ax\|_\infty. \tag{10.19}$$

Vu que $(Ax)_i = \sum_k A_{ik}x_k$, lorsque $\|x\|_\infty \leqslant 1$ nous avons $|(Ax)_i| \leqslant \sum_k |A_{ik}|$. Donc nous avons toujours

$$\|A\|_\infty \leqslant \max_i \sum_k |A_{ik}|. \tag{10.20}$$

△

Définition 10.7.

*La **topologie forte** sur l'espace des opérateurs est la topologie de la norme opérateur.*

Lorsque nous considérons un espace vectoriel d'applications linéaires, nous considérons toujours [1] dessus la topologie liée à cette norme.

1. Sauf lorsque les événements nous forceront à trahir.

Il existe aussi la **topologie faible** donnée par la notion de convergence[2] $A_i \to A$ si et seulement si $A_i x \to Ax$ pour tout $x \in E$.

Problèmes et choses à faire

Je crois, mais demande confirmation, que la topologie faible est celle des semi normes $\{p_v\}_{v\in E}$ données par $p_v(A) = \|A\|$. En effet la notion de convergence associée par la proposition 6.193 est $A_i \to A$ si et seulement si $p_v(A_i - A) \to 0$. Cette condition signifie $\|A_i(v) - A(v)\| \to 0$, c'est à dire $A_i(v) \to A(v)$.

Si le lecteur veut parler de cela au jury d'un concours, il est évident qu'il devra être capable d'ajouter des petits symboles au-dessus de toutes les flèches «→» du paragraphe précédent pour indiquer pour quelles topologies sont les convergences dont on parle.

Remarque 10.8.
Il faut noter que la topologie faible n'est pas une topologie métrique. Cela même si la condition $A_i x \to Ax$, elle, est métrique vu qu'elle est écrite dans E.

Dans le cas où E est de dimension infinie, la topologie faible est réellement différente de la topologie forte. Nous verrons à la sous-section 23.3.6 que dans le cas des projections sur un espaces de Hilbert, l'égalité

$$\sum_{i=1}^{\infty} \mathtt{proj}_{u_i} = \mathrm{Id} \tag{10.21}$$

est vraie pour la topologie faible, mais pas pour la topologie forte.

10.2.1 Norme d'algèbre

Définition 10.9 (Norme d'algèbre[121])**.**
Si A est une algèbre[3]*, une **norme d'algèbre** sur A est une norme telle que pour toute $u, v \in A$,*

$$\|uv\| \leqslant \|u\|\|v\|. \tag{10.22}$$

L'intérêt d'une norme d'algèbre est entre autres de mieux se comporter pour les séries, voir par exemple 15.12.

Définition 10.10 ([122])**.**
*Le **rayon spectral** d'une matrice carrée A, noté $\rho(A)$, est défini de la manière suivante :*

$$\rho(A) = \max_i |\lambda_i| \tag{10.23}$$

où les λ_i sont les valeurs propres de A.

10.11.
Quelque remarques sur la définition du rayon spectral.

- Même si A est une matrice réelle, les valeurs propres sont dans $\mathbb{C}$. Donc dans (10.23), $|\lambda_i|$ est le module dans $\mathbb{C}$ de λ_i.
- Vu que les valeurs propres de A sont les racines de son polynôme caractéristique (théorème 9.125), il y en a un nombre fini et le maximum est bien définit.
- La définition s'applique uniquement pour les espaces de dimension finie.

Lemme 10.12.
Soient des espaces vectoriels normés E et F, sur les corps $\mathbb{R}$ ou $\mathbb{C}$. Pour tout $A \in \mathcal{L}(E, F)$, et pour tout $u \in E$ nous avons la majoration

$$\|Au\| \leqslant \|A\|\|u\| \tag{10.24}$$

où la norme sur A est la norme opérateur subordonnée à la norme sur u.

Démonstration. Si $u \in E$ alors, étant donné que le supremum d'un ensemble est plus grand ou égal à chacun de éléments qui le compose,

$$\|A\| = \sup_{x\in E} \frac{\|Ax\|}{\|x\|} \geqslant \frac{\|Au\|}{\|u\|}, \tag{10.25}$$

donc le résultat annoncé : $\|Au\| \leqslant \|A\|\|u\|$. □

2. Est-ce qu'on peut décrire cette topologie à partir de ses ouverts ? Facilement ?
3. Définition 4.33.

Proposition 10.13 ([122]).
Soit une matrice $A \in \mathbb{M}(n,\mathbb{C})$ de rayon spectral $\rho(A)$. Soit une norme $\|.\|$ sur $\mathbb{C}^n$ et la norme opérateur correspondante. Alors

$$\rho(A) \leqslant \|A^k\|^{1/k} \tag{10.26}$$

pour tout $k \in \mathbb{N}$.

Démonstration. Soit $v \in \mathbb{C}^n$ et $\lambda \in \mathbb{C}$ un coupe vecteur-valeur propre. Nous avons $\|Av\| = |\lambda|\|v\|$ et aussi

$$|\lambda|^k\|v\| = \|\lambda^k v\| = \|A^k v\| \leqslant \|A^k\|\|v\|. \tag{10.27}$$

La dernière inégalité est due au fait que nous avons choisit sur $\mathbb{M}(n,\mathbb{C})$ la norme subordonnée à celle choisie sur $\mathbb{C}^n$, via le lemme 10.12. Nous simplifions par $\|v\|$ et obtenons $|\lambda| \leqslant \|A^k\|^{1/k}$. Étant donné que $\rho(A)$ est la maximum de tous les λ possibles, la majoration passe au maximum :

$$\rho(A) \leqslant \|A^k\|^{1/k}. \tag{10.28}$$

□

Lemme 10.14 (La norme opérateur est une norme d'algèbre[1]).
Soient des espaces vectoriels normés E, F et G. Soient des opérateurs linéaires bornés $B\colon E \to F$, $A\colon F \to G$. Alors

$$\|AB\| \leqslant \|A\|\|B\|. \tag{10.29}$$

Démonstration. Nous avons les (in)égalités suivantes :

$$\|AB\| = \sup_{x\in E} \frac{\|ABx\|_G}{\|x\|_E} \tag{10.30a}$$

$$= \sup_{\substack{x\in E\\ Bx\neq 0}} \frac{\|ABx\|}{\|x\|}\frac{\|Bx\|_F}{\|Bx\|_F} \tag{10.30b}$$

$$= \sup_{\substack{x\in E\\ Bx\neq 0}} \frac{\|ABx\|}{\|Bx\|}\frac{\|Bx\|}{\|x\|} \tag{10.30c}$$

$$\leqslant \underbrace{\sup_{\substack{x\in E\\ Bx\neq 0}} \frac{\|ABx\|}{\|Bx\|}}_{\leqslant\|A\|} \underbrace{\sup_{\substack{y\in E\\ By\neq 0}} \frac{\|Bx\|}{\|y\|}}_{=\|B\|} \tag{10.30d}$$

$$\leqslant \|A\|\|B\|. \tag{10.30e}$$

La dernière inégalité provient que dans $\sup_{\substack{x\in E\\ Bx\neq 0}} \|ABx\|/\|x\|$, le supremum est pris sur un ensemble plus petit que celui sur lequel porte la définition de la norme de A : seulement l'image de B au lieu de tout l'espace de départ de A. □

La chose impressionnante dans la proposition suivante est que $\rho(A)$ est définit indépendamment du choix de la norme sur $\mathbb{M}(n,\mathbb{K})$ ou sur $\mathbb{K}$. Lorsque nous écrivons $\|A\|$, nous disons implicitement qu'une norme a été choisie sur $\mathbb{K}$ et que nous avons pris la norme subordonnée sur $\mathbb{M}(n,\mathbb{K})$.

Proposition 10.15 ([123]).
Soit A une matrice de $\mathbb{M}(n,\mathbb{R})$ ou $\mathbb{M}(n,\mathbb{C})$. Alors

$$\rho(A) \leqslant \|A\|. \tag{10.31}$$

Démonstration. Nous devons séparer les cas suivant que le corps de base soit $\mathbb{R}$ ou $\mathbb{C}$.

Pour $A \in \mathbb{M}(n,\mathbb{C})$ Soit λ une valeur propre de A telle que $|\lambda|$ soit la plus grande. Nous avons donc $\rho(A) = |\lambda|$. Soit un vecteur propre $u \in \mathbb{C}^n$ pour la valeur propre λ. En prenant la norme sur l'égalité $Au = \lambda u$, et en utilisant le lemme 10.12,

$$|\lambda|\|u\| = \|Au\| \leqslant \|A\|\|u\|. \tag{10.32}$$

Donc $|\lambda| \leqslant \|A\|$ et $\rho(A) \leqslant \|A\|$.

Pour $A \in \mathbb{M}(n,\mathbb{R})$ L'endroit qui coince dans le raisonnement fait pour $\mathbb{M}(n,\mathbb{C})$ est que certes $A \in \mathbb{M}(n,\mathbb{R})$ possède une plus grande valeur propre en module et qu'un vecteur propre lui est associé. Mais ce vecteur propre est a priori dans $\mathbb{C}^n$, et non dans $\mathbb{R}^n$. Nous pouvons donc écrire $Au = \lambda u$, mais pas $\|Au\| = |\lambda|\|u\|$ parce que nous ne savons pas quelle norme prendre sur $\mathbb{C}^n$.

Il n'est pas certain que nous ayons une norme sur $\mathbb{C}^n$ qui se réduit sur $\mathbb{R}^n$ à celle choisie implicitement dans l'énoncé. Nous allons donc ruser un peu.

Soit une norme N sur $\mathbb{C}^n$ [4]. Nous nommons également N la norme subordonnée sur $\mathbb{M}(n,\mathbb{C})$ et la norme restreinte sur $\mathbb{M}(n,\mathbb{R})$. Vu que N est une norme sur $\mathbb{M}(n,\mathbb{R})$ et que ce dernier est de dimension finie, le théorème 10.3 nous indique que N est équivalente à $\|.\|$. Il existe donc $C > 0$ tel que

$$N(B) \leqslant C\|B\| \tag{10.33}$$

pour tout $B \in \mathbb{M}(n,\mathbb{R})$. Nous avons maintenant

$$\rho(A)^m \leqslant N(A^m) \leqslant C\|A^m\| \leqslant C\|A\|^m. \tag{10.34}$$

Justifications
— Par la proposition 10.13.
— Parce que $A^m \in \mathbb{M}(n,\mathbb{R})$ et la relation (10.33).
— Par itération du lemme 10.14.
Nous avons donc $\rho(A) \leqslant C^{1/m}\|A\|$ pour tout $m \in \mathbb{N}$. En prenant $m \to \infty$ et en tenant compte de $C^{1/m} \to 1$ nous trouvons $\rho(A) \leqslant \|A\|$.

□

Lemme 10.16 ([123])**.**
Soit $A \in \mathbb{M}(n,\mathbb{K})$ avec $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit $\epsilon > 0$. Il existe une norme algébrique sur $\mathbb{M}(n,\mathbb{K})$ telle que

$$N(A) \leqslant \rho(A) + \epsilon. \tag{10.35}$$

Démonstration. Soit par le lemme 9.152 une matrice inversible U telle que $T = UAU^{-1}$ soit triangulaire supérieure, avec les valeurs propres sur la diagonale. Notons que même si $A \in \mathbb{M}(n,\mathbb{R})$, les matrices U et T sont a priori complexes.

Soit $s \in \mathbb{R}$ ainsi que les matrices

$$D_s = \operatorname{diag}(1, s^{-1}, s^{-2}, \dots, s^{1-n}) \tag{10.36}$$

et $T_s = D_s T D_s^{-1}$. Nous fixerons un choix de s plus tard.

La norme que nous considérons est :

$$N(B) = \|(D_sU)B(D_sU)^{-1}\|_\infty \tag{10.37}$$

où $\|.\|_\infty$ est la norme sur $\mathbb{M}(,n\mathbb{K})$ subordonnée à la norme $\|.\|_\infty$ sur $\mathbb{K}^n$ dont nous avons déjà parlé dans l'exemple 10.6. Cela est bien une norme parce que
— Nous avons $\|B\|_\infty = 0$ si et seulement si $B = 0$, et vu que (D_sU) est inversible nous avons $(D_sU)B(D_sU)^{-1} = 0$ si et seulement si $B = 0$.
— $N(\lambda B) = |\lambda|N(B)$.
— Pour l'inégalité triangulaire :

$$\begin{aligned} N(B+C) &= \|(D_sU)B(D_sU)^{-1} + (D_sU)C(D_sU)^{-1}\|_\infty && \text{(10.38a)}\\ &\leqslant \|(D_sU)B(D_sU)^{-1}\|_\infty + \|(D_sU)C(D_sU)^{-1}\|_\infty && \text{(10.38b)}\\ &= N(B) + N(C). && \text{(10.38c)} \end{aligned}$$

4. Il y en a plein, par exemple celle du produit scalaire $\langle x, y\rangle = \sum_k x_k \bar{y}_k$.

En ce qui concerne la matrice A elle-même, nous avons

$$N(A) = \|(D_sU)A(D_sU)^{-1}\|_\infty = \|T_s\|_\infty. \tag{10.39}$$

C'est le moment de se demander comment se présente la matrice T_s. En tenant compte du fait que $(D_s)_{ik} = \delta_{ik}s^{1-i}$ nous avons

$$(T_s)_{ij} = \sum_{kl}(D_s)_{ik}T_{kl}(D_s^{-1})_{lj} = T_{ij}s^{j-i}. \tag{10.40}$$

La matrice T est encore triangulaire supérieure avec les valeurs propres de A sur la diagonale. Les éléments au-dessus de la diagonale sont tous multipliés par au moins s. Il est donc possible de choisir s suffisamment petit pour avoir [5]

$$\sum_{j=i+1}^{n} |(T_s)_{ij}| < \epsilon \tag{10.41}$$

Avec ce choix, la formule 10.20 donne

$$N(T_s) \leqslant \max_i \sum_k |(T_s)_{ik}| \leqslant \epsilon + \rho(A). \tag{10.42}$$

En effet le ϵ vient de la somme sur toute la ligne sauf la diagonale (c'est à dire la partie $k \neq i$) et du choix (10.41) pour s. Le $\rho(A)$ provient du dernier terme de la somme (le terme sur la diagonale) qui est une valeur propre de A, donc majorable par $\rho(A)$.

Nous devons encore prouver que N est une norme algébrique. Pour cela nous allons montrer qu'elle est subordonnée à la norme

$$\begin{aligned} n\colon \mathbb{K}^n &\to \mathbb{R}^+ \\ v &\mapsto \|(UD_s)v\|_\infty. \end{aligned} \tag{10.43}$$

Cela sera suffisant pour avoir une norme algébrique par le lemme 10.14. La norme n sur $\mathbb{K}^n$ produit la norme suivante sur $\mathbb{M}(n,\mathbb{K})$:

$$n(B) = \sup_{v\neq 0} \frac{n(B)}{n(v)} = \sup_{v\neq 0} \frac{\|(UD_s)Bv\|_\infty}{\|UD_sv\|_\infty}. \tag{10.44}$$

Vu que UD_s est inversible nous pouvons effectuer le changement de variables $v \mapsto (UD_s)^{-1}v$ pour écrire

$$n(B) = \sup_{v\neq 0} \frac{\|(UD_s)B(UD_s)^{-1}v\|_\infty}{\|(UD_s)(UD_s)^{-1}v\|_\infty} = \sup_{v\neq 0} \frac{\|(UD_s)B(UD_s)^{-1}v\|_\infty}{\|v\|_\infty} = \|(UD_s)B(UD_s)^{-1}\|_\infty = N(B). \tag{10.45}$$

□

Proposition 10.17.
Si $A \in \mathbb{M}(n,\mathbb{R})$ alors $\rho(A)^m = \rho(A^m)$ pour tout $m \in \mathbb{N}$.

Démonstration. La matrice A peut être vue dans $\mathbb{M}(n,\mathbb{C})$ et nous pouvons lui appliquer le corollaire 9.158 :

$$\mathrm{Spec}(A^k) = \{\lambda^k \text{ tel que } \lambda \in \mathrm{Spec}(A)\}. \tag{10.46}$$

À noter qu'il n'y a pas de magie : le spectre de la matrice réelle A est déjà défini en voyant A comme matrice complexe. Le spectre dont il est question dans (10.46) est bien celui dont on parle dans la définition du rayon spectral.

5. Il me semble qu'il manque un module dans [123].

Nous avons ensuite :

$$\rho(A^k) = \max\{|\lambda| \text{ tel que } \lambda \in \text{Spec}(A^k)\} \tag{10.47a}$$
$$= \max\{|\lambda^k| \text{ tel que } \lambda \in \text{Spec}(A)\} \tag{10.47b}$$
$$= \max\{|\lambda|^k \text{ tel que } \lambda \in \text{Spec}(A)\} \tag{10.47c}$$
$$= \rho(A)^k. \tag{10.47d}$$

□

Proposition 10.18 (Bornée si et seulement si continue[124])**.**
Soient E et F des espaces vectoriels normés. Une application linéaire $E \to F$ est bornée si et seulement si elle est continue.

Démonstration. Nous commençons par supposer que A est bornée. Par le lemme 10.14, pour tout $x, y \in E$, nous avons

$$\|A(x) - A(y)\| = \|A(x - y)\| \leqslant \|A\|\|x - y\|. \tag{10.48}$$

En particulier si $x_n \xrightarrow{E} x$ alors

$$0 \leqslant \|A(x_n) - A(x)\| \leqslant \|A\|\|x - x_n\| \to 0 \tag{10.49}$$

et A est continue en vertu de la caractérisation séquentielle de la continuité, proposition 6.138.

Supposons maintenant que $\|A\|$ ne soit pas borné, c'est à dire que l'ensemble $\{\|A(x)\| \text{ tel que } \|x\| = 1\}$ ne soit pas borné. Alors pour tout $k \geqslant 1$ il existe $x_k \in B(0, 1)$ tel que $\|A(x_k)\| > k$. La suite x_k/k tend vers zéro parce que $\|x_k\| = 1$, mais $\|A(x_k)\| \geqslant 1$ pour tout k. Cela montre que A n'est pas continue. □

Définition 10.19 ([125])**.**
Soient E et F deux espaces vectoriels normés.

- *L'ensemble des applications linéaires $E \to F$ est noté $\mathcal{L}(E, F)$.*
- *Un **morphisme** est une application linéaire $E \to F$ continue pour la topologie de la norme opérateur. Nous avons vu dans la proposition 10.18 que la continuité était équivalente à être bornée. L'ensemble des morphismes est noté $L(E, F)$.*
- *Un **isomorphisme** est un morphisme continu inversible dont l'inverse est continu. Nous notons $\text{GL}(E, F)$ l'ensemble des isomorphismes entre E et F.*

Le point important de la définition 10.19 est la continuité. En dimension infine, la continuité n'est par exemple pas équivalente à l'inversibilité (penser à $e_k \mapsto ke_k$).

10.2.2 Application linéaire continue et bornée

Nous avons vu dans la proposition 10.18 que pour une application linéaire, être bornée est équivalent à être continue. Nous allons maintenant voir un certain nombre d'exemples illustrant ce fait.

Exemple 10.20(Une application linéaire non continue)
Soit V l'espace vectoriel normé des suites *finies* de réels muni de la norme usuelle $\|c\| = \sqrt{\sum_{i=0}^{\infty} |c_i|^2}$ où la somme est finie. Nous nommons $\{e_k\}_{k\in\mathbb{N}}$ la base usuelle de cet espace, et nous considérons l'opérateur $f : V \to V$ donnée par $f(e_k) = ke_k$. C'est évidemment linéaire, mais ce n'est pas continu en zéro. En effet la suite $u_k = e_k/k$ converge vers 0 alors que $f(u_k) = e_k$ ne converge pas. △

Cet exemple aurait pu également être donnée dans un espace de Hilbert, mais il aurait fallu parler de domaine.

Exemple 10.21(Une autre application linéaire non continue[126])
En dimension infinie, une application linéaire n'est pas toujours continue. Soit E l'espace des

polynômes à coefficients réels sur $[0,1]$ muni de la norme uniforme. L'application de dérivation $\varphi\colon E \to E$, $\varphi(P) = P'$ n'est pas continue.

Pour la voir nous considérons la suite $P_n = \frac{1}{n}X^n$. D'une part nous avons $P_n \to 0$ dans E parce que $P_n(x) = \frac{x^n}{n}$ avec $x \in [0,1]$. Mais en même temps nous avons $\varphi(P_n) = X^{n-1}$ et donc $\|\varphi(P_n)\| = 1$.

Nous n'avons donc pas $\lim_{n\to\infty} \varphi(P_n) = \varphi(\lim_{n\to\infty} P_n)$ et l'application φ n'est pas continue en 0. Elle n'est donc continue nulle part par linéarité.

Nous avons utilisé le critère séquentiel de la continuité, voir la définition 6.137 et la proposition 6.138. △

Remarque 10.22.
Cette proposition permet de retrouver l'exemple 10.20 plus simplement. Si $\{e_k\}_{k\in\mathbb{N}}$ est une base d'un espace vectoriel normé formée de vecteurs de norme 1, alors l'opérateur linéaire donné par $u(e_k) = ke_k$ n'est pas borné et donc pas continu.

C'est également ce résultat qui montre que le produit scalaire est continu sur un espace de Hilbert par exemple.

Lemme 10.23.
Soit F un espace de Banach et deux suites $A_k \to A$ et $B_k \to B$ dans $\mathcal{L}(F,F)$. Alors $A_k \circ B_k \to A \circ B$ dans $\mathcal{L}(F,F)$, c'est à dire

$$\lim_{n\to\infty}(A_kB_k) = \left(\lim_{n\to\infty} A_k\right)\left(\lim_{n\to\infty} B_k\right). \tag{10.50}$$

Démonstration. Il suffit d'écrire

$$\|A_kB_k - AB\| \leqslant \|A_kB_k - A_kB\| + \|A_kB - AB\|. \tag{10.51}$$

Le premier terme tend vers zéro pour $k \to \infty$ parce que

$$\begin{aligned}\|A_kB_k - A_kB\| &= \|A_k(B_k - B)\| && \text{(10.52a)}\\ &\leqslant \|A_k\|\|B_k - B\| \to \|A\| \cdot 0 && \text{(10.52b)}\\ &= 0 && \text{(10.52c)}\end{aligned}$$

où nous avons utilisé la propriété fondamentale de la norme opérateur : la proposition 10.18. Le second terme tend également vers zéro pour la même raison. □

Proposition 10.24 (Distributivité de la somme infinie)**.**
Soient E un espace normé, une suite (u_k) dans $\mathrm{GL}(E)$ ainsi que $a \in \mathrm{GL}(E)$. Pourvu que la série $\sum_{n=0}^{\infty} u_k$ converge nous avons

$$\left(\sum_{k=0}^{\infty} u_k\right) a = \sum_{k=0}^{\infty}(u_k a). \tag{10.53}$$

Démonstration. Par définition de la somme infinie,

$$\spadesuit = \left(\sum_{k=0}^{\infty} u_k\right) a = \left(\lim_{n\to\infty}\sum_{k=0}^{n} u_k\right) a. \tag{10.54}$$

Le lemme 10.23 appliqué à $n \mapsto \sum_{k=0}^{n} u_k$ et à la suite constante a nous donne

$$\spadesuit = \lim_{n\to\infty}\left(\sum_{k=0} u_k a\right), \tag{10.55}$$

ce que nous voulions par distributivité de la somme finie : dans (10.55), le a est dans ou hors de la somme, au choix. L'important est qu'il soit dans la limite. □

10.2.3 Normes de matrices et d'applications linéaires

Théorème 10.25 (Norme matricielle et rayon spectral[127])**.**
La norme 2 *d'une matrice est liée au rayon spectral de la façon suivante :*

$$\|A\|_2 = \sqrt{\rho(A^tA)} \tag{10.56}$$

ou plus généralement par $\|A\|_2 = \sqrt{\rho(A^*A)}$.

Lemme 10.26.
Soit une matrice $A \in \mathbb{M}(n,\mathbb{R})$ *qui est symétrique, strictement définie positive. Soient* λ_{min} *et* λ_{max} *les plus petites et plus grandes valeurs propres. Alors*

$$\|A\|_2 = \lambda_{max} \qquad et \qquad \|A^{-1}\|_2 = \frac{1}{\lambda_{min}}. \tag{10.57a}$$

Démonstration. Soient les vecteurs $v_1,\dots,v_n$ formant une base orthonormée de vecteurs propres[6] de A. Nous notons v_{max} celui de λ_{max}. Nous avons :

$$\|A\|_2 \geqslant \|Av_{max}\| = |\lambda_{max}|\|v_{max}\| = |\lambda_{max}| = \lambda_{max}. \tag{10.58}$$

Voila l'inégalité dans un sens. Montrons l'inégalité dans l'autre sens. Soit $x = \sum_i x_i v_i$ avec $\|x\|_2 = 1$. Alors

$$\|Ax\| = \|\sum_i x_i\lambda_i v_i\| \leqslant \sqrt{\sum_i x_i^2\lambda_i^2} \leqslant \lambda_{max}\sqrt{\sum_i x_i^2} = \lambda_{max}. \tag{10.59}$$

En ce qui concerne l'affirmation pour la norme de A^{-1}, il suffit de remarquer que ses valeurs propres sont les inverses des valeurs propres de A. □

Proposition 10.27.
La fonction

$$\begin{aligned} f\colon \mathbb{M}(n,\mathbb{R})\times\mathbb{M}(n,\mathbb{R}) &\to \mathbb{R} \\ (X,Y) &\mapsto \mathrm{Tr}(X^tY) \end{aligned} \tag{10.60}$$

est un produit scalaire sur $\mathbb{M}(n,\mathbb{R})$.

Démonstration. Il faut vérifier la définition 9.24.

- La bilinéarité est la linéarité de la trace.
- La symétrie de f est le fait que $\mathrm{Tr}(A^t) = \mathrm{Tr}(A)$.
- L'application f est définie positive parce que si $X \in \mathbb{M}$, alors X^tX est symétrique définie positive, donc diagonalisable avec des nombres positifs sur la diagonale. La trace étant un invariant de similitude, nous avons $f(X,X) = \mathrm{Tr}(X^tX) \geqslant 0$. De plus si $\mathrm{Tr}(X^tX) = 0$, alors $X^tX = 0$ (pour la même raison de diagonalisation). Mais alors $\|Xu\| = 0$ pour tout $u \in E$, ce qui signifie que $X = 0$.

□

Exemple 10.28
Soit $m = n$, un point λ dans $\mathbb{R}$ et T_λ l'application linéaire définie par $T_\lambda(x) = \lambda x$. La norme de T_λ est alors

$$\|T_\lambda\|_{\mathcal{L}} = \sup_{\|x\|_{\mathbb{R}^m}\leqslant 1} \|\lambda x\|_{\mathbb{R}^n} = |\lambda|.$$

Notez que T_λ n'est rien d'autre que l'homothétie de rapport λ dans $\mathbb{R}^m$. △

6. Possible par le théorème spectral 9.165.

Exemple 10.29
Considérons la rotation T_α d'angle α dans $\mathbb{R}^2$. Elle est donnée par l'équation matricielle

$$T_\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\alpha)x + \sin(\alpha)y \\ -\sin(\alpha)x + \cos(\alpha)y \end{pmatrix} \tag{10.61}$$

Étant donné que cela est une rotation, c'est une isométrie : $\|T_\alpha x\| = \|x\|$. En ce qui concerne la norme de T_α nous avons

$$\|T_\alpha\| = \sup_{x\in\mathbb{R}^2} \frac{\|T_\alpha(x)\|}{\|x\|} = \sup_{x\in\mathbb{R}^2} \frac{\|x\|}{\|x\|} = 1. \tag{10.62}$$

Toutes les rotations dans le plan ont donc une norme 1. La même preuve tient pour toutes les rotations en dimension quelconque. △

Exemple 10.30
Soit $m = n$, un point b dans $\mathbb{R}^m$ et T_b l'application linéaire définie par $T_b(x) = b \cdot x$ (petit exercice : vérifiez qu'il s'agit vraiment d'une application linéaire). La norme de T_b satisfait les inégalités suivantes

$$\|T_b\|_{\mathcal{L}} = \sup_{\|x\|_{\mathbb{R}^m}\leqslant 1} \|b \cdot x\|_{\mathbb{R}^n} \leqslant \sup_{\|x\|_{\mathbb{R}^m}\leqslant 1} \|b\|_{\mathbb{R}^n}\|x \cdot x\|_{\mathbb{R}^n} \leqslant \|b\|_{\mathbb{R}^n},$$

$$\|T_b\|_{\mathcal{L}} = \sup_{\|x\|_{\mathbb{R}^m}\leqslant 1} \|b \cdot x\|_{\mathbb{R}^n} \geqslant \left\| b \cdot \frac{b}{\|b\|_{\mathbb{R}^n}} \right\|_{\mathbb{R}^n} = \|b\|_{\mathbb{R}^n},$$

donc $\|T_b\|_{\mathcal{L}} = \|b\|_{\mathbb{R}^n}$. △

Proposition 10.31.
Une application linéaire de $\mathbb{R}^m$ dans $\mathbb{R}^n$ est continue.

Démonstration. Soit x un point dans $\mathbb{R}^m$. Nous devons vérifier l'égalité

$$\lim_{h\to 0_m} T(x+h) = T(x). \tag{10.63}$$

Cela revient à prouver que $\lim_{h\to 0_m} T(h) = 0$, parce que $T(x+h) = T(x) + T(h)$. Nous pouvons toujours majorer $\|T(h)\|_n$ par $\|T\|_{\mathcal{L}(\mathbb{R}^m,\mathbb{R}^n)}\|h\|_{\mathbb{R}^m}$ (lemme 10.12). Quand h s'approche de 0_m sa norme $\|h\|_m$ tend vers 0, ce que nous permet de conclure parce que nous savons que de toutes façons, $\|T\|_{\mathcal{L}}$ est fini. □

Note : dans un espace de dimension infinie, la linéarité ne suffit pas pour avoir la continuité : il faut de plus être borné (ce que sont toutes les applications linéaires $\mathbb{R}^m \to \mathbb{R}^n$). Voir la proposition 10.18.

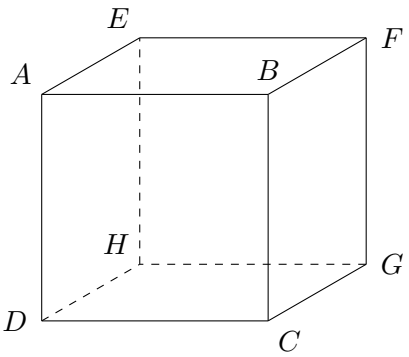
10.3 Isométriques du cube

Les isométries du cube proviennent de [2].

Nous considérons le cube centré en l'origine de $\mathbb{R}^3$ et G, le groupe des isométries de $\mathbb{R}^3$ préservant ce cube. Nous notons aussi G^+ le sous-groupes de G constitué des éléments de déterminant positif. Nous notons

$$\mathcal{D} = \{D_1, \dots, D_4\} \tag{10.64}$$

l'ensemble des grandes diagonales, c'est à dire les segments $[AG]$, $[EC]$, $[FD]$, et $[BH]$. Nous savons que G préserve les longueurs et que ces segments sont les plus longs possibles à l'intérieur du



cube. Donc G agit sur $\mathcal{D}$ parce qu'il ne peut transformer une grande diagonale qu'en une autre grande diagonale. Nous avons donc un morphisme de groupes

$$\rho\colon G \to S_4. \tag{10.65}$$

Nous montrons ce que morphisme est surjectif en montrant qu'il contient les transpositions. Le groupe G contient la symétrie axiale passant par le milieu M de $[AE]$ et le milieu N de CG. Il est facile de voir que cette symétrie permute $[AG]$ avec $[EC]$. De plus elle laisse $[FD]$ inchangée. En effet, aussi incroyable que cela paraisse en regardant le dessin, nous avons $FD \perp MN$, parce qu'en termes de vecteurs directeurs,

$$\overrightarrow{ON} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \overrightarrow{OF} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}. \tag{10.66}$$

Étudions à présent le noyau $\ker(\rho)$. Si $f \in \ker(\rho)$ n'est pas l'identité, alors $f(D_i) = D_i$ pour tout i, mais au moins pour une des diagonales les sommets sont inversés. Quitte à renommer les sommets du cube nous supposons que la diagonale $[AG]$ est retournée : $f(A) = G$ et $f(G) = A$. Regardons où peut partir B sous l'effet de f. Étant donné que f préserve les diagonales, $f(B) \in \{B, C\}$, mais étant donné que f est une isométrie, $d\big(f(B), f(G)\big) = d(B, G)$, et nous concluons que $f(B) = H$. Donc la diagonale $[BH]$ est retournée sous l'effet de f. En raisonnant de même, nous voyons que f retourne toutes les diagonales. Donc les éléments non triviaux de $\ker(\rho)$ retournent toutes les diagonales ; il n'y en a donc qu'un seul et c'est la symétrie centrale :

$$\ker(\rho) = \{\mathrm{Id}, s_0\}. \tag{10.67}$$

Le premier théorème d'isomorphisme 3.22 nous permet d'écrire le quotient de groupes :

$$\frac{G}{\{\mathrm{Id}, s_0\}} \simeq S_4. \tag{10.68}$$

Une classe d'équivalence modulo $\ker(\rho)$ dans G est donc toujours de la forme $\{f, f \circ s_0\}$. Et vu que s_0 est de déterminant -1, une classe contient toujours exactement un élément de déterminant 1 et un de déterminant -1.

D'autre part $\ker(\rho)$ est normal dans G parce que en tant que matrice, $s_0 = -\mathbb{1}$, donc les problèmes de commutativité ne se posent pas. L'application

$$\begin{aligned} \varphi\colon \frac{G}{\{\mathrm{Id}, s_0\}} &\to G^+ \\ [g] &\mapsto \begin{cases} g & \text{si } \det(g) > 0 \\ g \circ s_0 & \text{sinon} \end{cases} \end{aligned} \tag{10.69}$$

est un isomorphisme de groupes. Et enfin nous pouvons écrire

$$G^+ \simeq S_4. \tag{10.70}$$

Nous allons maintenant utiliser le corollaire 3.95 pour montrer que $G = G^+ \times_\sigma \ker(\rho)$. Les conditions sont remplies :

— $\ker(\rho)$ normalise G^+ parce que $\ker(\rho)$ ne contient que $\pm\mathbb{1}$.
— $\ker(\rho) \cap G^+ = \{\mathrm{Id}\}$.
— $\ker(\rho)G^+ = G$ parce que les classes d'équivalence de G modulo $\ker(\rho)$ sont composées de $\{f, f \circ s_0\}$.

Vu que $G^+ \simeq S_4$ et $\ker(\rho) \simeq \mathbb{Z}/2\mathbb{Z}$ nous pouvons écrire de façon plus brillante que

$$G \simeq S_4 \times_\sigma \mathbb{Z}/2\mathbb{Z}. \tag{10.71}$$

Mais étant donné que la conjugaison par s_0 est triviale, le produit semi-direct est un produit direct :

$$G \simeq S_4 \times \mathbb{Z}/2\mathbb{Z}. \tag{10.72}$$

Il est maintenant du meilleur goût de pouvoir identifier géométriquement ces éléments. Les éléments de $\mathbb{Z}/2\mathbb{Z} = \{\mathrm{Id}, s_0\}$ ne font pas de mystères. Dans S_4 nous avons les classes de conjugaison des éléments Id, (12), (123), (1234) et (12)(34) déterminées durant l'exemple 8.36.

(1) L'élément (12) consiste à permuter deux diagonales et laisser les autres en place. Nous avons déjà vu que c'était une symétrie axiale passant par les milieux de deux côtés opposés. Cela fait 6 axes d'ordre 2.

(2) L'élément (123) fixe une des diagonales. C'est donc la symétrie axiale le long de la diagonale fixée. Par exemple la symétrie d'axe (AG) fait bouger le point B de la façon suivante :

$$B \to D \to E \to B. \tag{10.73}$$

C'est une rotation est d'angle $\frac{2\pi}{3}$. Cela sont 8 rotations d'ordre 3.

Notons à ce propos que la différence entre (234) et (243) est que la première fait une rotation d'angle $2\pi/3$ tandis que la seconde fait une rotation d'angle $-2\pi/3$.

(3) L'élément (1234) ne maintient aucune des diagonales et est d'ordre 4. C'est donc la rotation d'angle $\pi/2$ ou $-\pi/2$ autour de l'axe passant par les milieux de deux faces opposées. Il y en a 6 comme ça (3 paires de faces puis pour chaque il y a $\pi/2$ et $-\pi/2$), et ça tombe bien 6 est justement la taille de la classe de conjugaison de (1234) dans S_4.

(4) L'élément (12)(34) est le carré de la précédente[7], c'est à dire les rotations d'angle π autour des mêmes axes. Cela fait 3 éléments d'ordre 2.

10.4 Chiffrement RSA

Ce passage sur RSA provient en bonne partie de la la page Wikipédia.

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.

10.4.1 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 8.30, 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)[8].

Bob envoie la paire (n, e) à Alice.

Remarque 10.32.
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.

10.4.2 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

$$C = M^e \mod n. \tag{10.74}$$

7. En fait c'est (13)(24), le carré de la précédente, mais c'est la même classe de conjugaison.
8. Le fait que e soit public et d soit privé est une convention. e comme *encryption* et d comme *decryption*.

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.

10.4.3 Déchiffrement

Nous allons montrer que $M = C^d \mod n$, et donc que Bob, connaissant (n, d), peut déchiffrer. D'abord

$$C^d = (M^e)^d = M^{ed}, \tag{10.75}$$

mais nous savons qu'il existe k tel que

$$ed = 1 + k\varphi(n) = 1 + k(p-1)(q-1). \tag{10.76}$$

L'étape astucieuse est de remarquer que

$$M^{1+k(p-1)(q-1)} \in [M]_p \cap [M]_q. \tag{10.77}$$

Pour montrer cela nous utilisons le petit théorème de Fermat 5.13 et la remarque 5.14.

— 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 (10.77). Le nombre $M^{1+k\varphi(n)} - M$ est donc à la fois multiple de p et de q.

Le lemme chinois 5.23 nous dit immédiatement [9] qu'alors

$$M^{1+k\varphi(n)} - M \tag{10.78}$$

est un multiple de $pq = n$, c'est à dire que

$$C^d = M^{ed} \in [M]_n. \tag{10.79}$$

Si on ne croit pas au lemme chinois, on peut utiliser le lemme de Gauss. Posons

$$M^{1+k\varphi(n)} - M = ap = bq. \tag{10.80}$$

Dans ce cas p divise bq, mais q est premier avec p, donc le lemme de Gauss 4.78 nous enseigne [10] que p divise b.

10.4.4 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 si 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 ?

9. 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.
10. Ici aussi, si $p = q$, ça ne marche pas.

10.4.5 Problèmes calculatoires

Pour implémenter RSA, il faut pouvoir faire (au moins) trois choses :

(1) Trouver de grands nombres premiers.

(2) Trouver des couples de Bézout.

(3) Calculer M^e lorsque e est très grand.

En ce qui concerne le problème de trouver des nombres premiers, c'est compliqué, mais il faut savoir qu'il y en a plein. À 120 chiffres, il y a environ autant de nombres premiers que d'atomes dans 10^{20} fois l'univers connu. Cela rend impossible toute tentative de factoriser un grand nombre en essayant toutes les possibilités. Même pas en science-fiction [11].

Trouver des nombres u et v tels que $Au + Bv = \text{pgcd}(A, B)$ est un problème expliqué en 3.5.4.

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

$$M^e = (M^k)^2; \tag{10.81}$$

si $e = 2k + 1$ alors nous calculons

$$M^e = M(M^k)^2. \tag{10.82}$$

Le calcul prend alors seulement environ $\log_2(e)$ étapes. Pour donner une idée,

$$\log_2(10^{120}) \simeq 400. \tag{10.83}$$

Très raisonnable, mais un ordinateur reste indispensable.

10.4.6 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.

10.4.7 Note non mathématique pour doucher l'enthousiasme

Il est souvent dit[128] que différents systèmes de chiffrement peuvent aider à avoir des discussions «discrètes» dans les régimes totalitaires. La technologie au service de la démocratie, voila qui enthousiasme la jeunesse [12]. 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.

10.5 Suites

10.5.1 Limites, convergence

Nous disons qu'une suite réelle (x_n) converge [13] vers ℓ lorsque pour tout ε, il existe un M tel que

$$n > N \Rightarrow |x_n - \ell| \leqslant \varepsilon. \tag{10.84}$$

11. Cela donne une idée des connaissances en math des klingons, dont le docteur Spock parvient à craquer le code mentalement en deux heures.

12. Cela dit, le navigateur Tor[129], qui est un pur produit de RSA, permet effectivement d'accéder en France aux sites bloqués pour apologie du terrorisme (mars 2015).

13. Voir la définition 6.90 pour plus de détail.

Le concept fondamental de cette définition est la notion de valeur absolue qui permet de donner la «distance» entre deux réels. Dans un espace vectoriel normé quelconque, cette notion est généralisée par la distance associée à la norme (définition 6.33). Nous pouvons donc facilement définir le concept de convergence d'une suite dans un espace vectoriel normé.

Définition 10.33.
Soit une suite (x_n) dans un espace vectoriel normé V. Nous disons qu'elle est ***convergente*** *s'il existe un élément $\ell \in V$ tel que*

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ tel que } n \geqslant N \Rightarrow \|x_n - l\| < \varepsilon. \tag{10.85}$$

Dans ce cas, ℓ est appelé la ***limite*** *de la suite (x_n).*

Lemme 10.34.
Soit (x_n) une suite convergente contenue dans un ensemble $A \subset V$. Alors la limite x_n appartient à $\bar{A}$.

Démonstration. Supposons que nous ayons une partie A de V, et une suite (x_n) dont la limite ℓ se trouve hors de $\bar{A}$. Dans ce cas, il existe un $r > 0$ tel que [14] $B(\ell, r) \cap A = \varnothing$. Si tous les éléments x_n de la suite sont dans A, il n'y en a donc aucun tel que $d(x_n, \ell) = \|x_n - \ell\| < r$. Cela contredit la notion de convergence $x_n \to \ell$. □

Nous avons déjà mentionné dans l'exemple 9.78 que zéro était un point adhérent à l'ensemble $F = \{(-1)^n/n \text{ tel que } n \in \mathbb{N}_0\}$. Nous savons maintenant que 0 étant la limite de la suite, il est automatiquement adhérent à l'ensemble des éléments de la suite.

Corollaire 10.35.
Soit a un point de l'adhérence d'une partie A de V. Alors il existe une suite d'éléments dans A qui converge vers a.

Démonstration. Si $a \in A$, alors nous pouvons prendre la suite constante $x_n = a$. Si a n'est pas dans A, alors a est dans ∂A, et pour tout n, il existe un point de A dans la boule $B(a, \frac{1}{n})$. Si nous nommons x_n ce point, la suite ainsi construite est une suite contenue dans A et qui converge vers a (ce dernier point est laissé à la sagacité du lecteur ou de la lectrice). □

En termes savants, ce corollaire signifie que la fermeture $\bar{A}$ est composé de A plus de toutes les limites de toutes les suites contenues dans A.

10.5.2 Critère de Cauchy

Lemme 10.36.
Une suite de Cauchy dans un espace vectoriel normé admettant une sous-suite convergente est elle-même convergente vers la même limite.

Démonstration. Soit (a_n) une suite de Cauchy dans un espace vectoriel normé E et ℓ la limite d'une sous-suite de (a_n). Soit $\epsilon > 0$ et $N \in \mathbb{N}$ tel que $\|a_m - a_p\| < \epsilon$ dès que $m, p \geqslant N$. Nous allons montrer que si $k > N$ alors $\|a_k - \ell\| < 2\epsilon$. Pour cela nous considérons un $n > N$ tel que $\|a_n - \ell\| \leqslant \epsilon$ et nous calculons

$$\|a_k - \ell\| \leqslant \|a_k - a_n\| + \|a_n - \ell\| \leqslant 2\epsilon. \tag{10.86}$$

□

Théorème 10.37 (Critère de Cauchy)**.**
Une suite réelle est convergence si et seulement si elle est de Cauchy. En particulier $\mathbb{R}$ est un espace complet.

14. Une autre manière de dire la même chose : si $\ell \notin \bar{A}$, alors $d(\ell, A) > 0$.

Démonstration. Soit (a_n) une suite de Cauchy dans $\mathbb{R}$. □

Définition 10.38.
*Nous disons que deux suites (u_n) et (v_n) sont **équivalentes** s'il existe une fonction $\alpha\colon \mathbb{N} \to \mathbb{R}$ telle que*

(1) pour tout n à partir d'un certain rang, $u_n = v_n\alpha(n)$

(2) $\alpha(n) \to 1$.

Lemme 10.39.
Si les suites (u_n) et (v_n) sont équivalentes et si (v_n) admet une limite l différente de 1, alors les suites $(\ln u_n)$ et $(\ln v_n)$ sont équivalentes.

Démonstration. En effet si $u_n = v_n\alpha(n)$ alors

$$\ln(u_n) = \ln(v_n) + \ln\big(\alpha(n)\big) = \ln(v_n)\left(1 + \frac{\ln\big(\alpha(n)\big)}{\ln(v_n)}\right), \tag{10.87}$$

et comme $\alpha(n) \to 1$, la parenthèse tend vers 1. □

Lemme 10.40 (Formule de Stirling[130])**.**
Nous avons l'équivalence de suites

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}. \tag{10.88}$$

10.5.3 Approximation

Le lemme suivant est surtout intéressant en dimension infinie.

Lemme 10.41.
Soit un espace vectoriel normé V et un sous-espace vectoriel dense A. Soit $v \in V$; il existe une suite (v_n) dans A telle que $v_n \xrightarrow{V} v$ et $\|v_n\| \leqslant \|v\|$ pour tout n.

Démonstration. Vu que A est dense, il existe une suite a_n dans A telle que $a_n \to v$. Ensuite il suffit de poser

$$v_n = \frac{n}{n+1}\frac{\|v\|}{\|a_n\|}a_n. \tag{10.89}$$

Par construction nous avons toujours

$$\|v_n\| = \frac{n}{n+1}\|v\| \leqslant \|v\|. \tag{10.90}$$

Et de plus, la norme étant continue[15],

$$\lim_{n\to\infty} v_n = \lim_{n\to\infty} \frac{n}{n+1} \lim_{n\to\infty} \frac{\|v\|}{\|v_n\|} \lim_{n\to\infty} v_n = v. \tag{10.91}$$

Le fait que v_n soit dans A est dû au fait que A soit vectoriel. □

Proposition 10.42.
Soit un espace vectoriel normé V et un sous-espace vectoriel dense A. Soit $v \in V$; pour tout $a \in \mathbb{R}$ nous avons

$$\sup\{|v \cdot a| \text{ tel que } a \in A \text{ et } \|a\| \leqslant \lambda\} = \lambda\|v\|. \tag{10.92}$$

15. Où dans le calcul suivant nous utilisons la continuité de la norme ? Posez-vous la question.

Démonstration. D'abord pour tout $a \in A$ vérifiant $\|a\| \leqslant \lambda$ l'inégalité de Cauchy-Schwarz 9.25 donne

$$|v \cdot a| \leqslant \|v\|\|a\| \leqslant \lambda\|v\|. \tag{10.93}$$

Donc le supremum dont on parle est majoré par $\lambda\|v\|$.

Il nous faut l'inégalité dans l'autre sens. Par densité nous pouvons choisir une suite $v_n \in A$ tel que $v_n \to v$. Ensuite nous posons

$$a_n = \frac{\lambda}{\|v_n\|}v_n. \tag{10.94}$$

Nous avons $\|a_n\| = \lambda$ pour tout n et

$$|v \cdot a_n| = \frac{\lambda}{\|v_n\|}|v \cdot v_n|, \tag{10.95}$$

et en passant à la limite,

$$\lim_{n\to\infty} |v \cdot a_n| = \frac{\lambda}{\|v\|}\|v \cdot v\| = \lambda\|v\|. \tag{10.96}$$

Donc l'ensemble sur lequel nous prenons le supremum contient une suite convergente vers $\lambda\|v\|$. Le supremum est donc au moins aussi grand que cela. □

10.6 Séries

Définition 10.43.
*Soit (a_k) une suite dans un espace vectoriel normé $(V, \|.\|)$. La suite des **sommes partielles** associée est la suite (s_k) définie par*

$$s_k = \sum_{i=0}^{k} a_i \tag{10.97}$$

*La **série** associée est la limite des sommes partielles*

$$\sum_{n=0}^{\infty} a_k = \lim_{k\to\infty} \sum_{k=0}^{n} a_k \tag{10.98}$$

si elle existe.

*Si une telle limite existe nous disons que $\sum_{k=0}^{\infty} a_k$ **converge** dans V. Si la limite de la suite des sommes partielles n'existe pas nous disons que la série **diverge**.*

Remarque 10.44.
Si la limite de la suite des sommes partielles n'existe pas dans V, alors elle peut parfois exister dans des extensions de V. Par exemple une série de rationnels convergeant vers $\sqrt{2}$ dans $\mathbb{R}$ ne converge pas dans $\mathbb{Q}$. Autre exemple : avec une bonne topologie sur $\bar{\mathbb{R}}$, une série peut ne pas converger dans $\mathbb{R}$ mais converger vers $\pm\infty$ dans $\bar{\mathbb{R}}$.

Dans le cas des espaces de fonctions, nous avons une norme importante : la norme uniforme définie par $\|f\|_\infty = \sup\{f(x)\}$ où le supremum est pris sur l'ensemble de définition de f.

Définition 10.45 (Convergence absolue)**.**
*Nous disons que la série $\sum_{n=0}^{\infty} a_n$ dans l'espace vectoriel normé V **converge absolument** si la série $\sum_{n=0}^{\infty} \|a_n\|$ converge dans $\mathbb{R}$.*

La convergence absolue dans le cas d'un espace de fonctions muni de la norme uniforme s'appelle la convergence normale.

Définition 10.46 (Convergence normale)**.**
*Une série de fonctions $\sum_{n\in\mathbb{N}} u_n$ converge **normalement** si la série de nombres $\sum_n \|u_n\|_\infty$ converge. C'est à dire si la série converge absolument pour la norme $\|f\|_\infty$.*

La convergence normale est à ne pas confondre avec la convergence uniforme.

Définition 10.47 (Convergence uniforme)**.**
*La somme $\sum_n f_n$ **converge uniformément** vers la fonction F si la suite des sommes partielles converge uniformément, c'est à dire si*

$$\lim_{N\to\infty} \| \sum_{n=1}^{N} f_n - F\|_\infty = 0. \tag{10.99}$$

Proposition 10.48.
Une série convergeant absolument dans un espace de Banach[16] *y converge au sens usuel.*

Démonstration. Soit (a_k) une suite dans un espace vectoriel normé complet dont la série converge absolument. Nous allons montrer que la suite des sommes partielles est de Cauchy. Cela suffira à montrer sa convergence par hypothèse de complétude.

Nous avons

$$\|s_p - s_l\| = \| \sum_{k=l+1}^{p} a_k\| \leqslant \sum_{k=l+1}^{p} \|a_k\| = \bar{s}_p - \bar{s}_l \tag{10.100}$$

où $\bar{s}_n = \sum_{k=0}^{n} \|a_k\|$ est la suite des sommes partielles de la série des normes (qui converge). Vu que la suite $(\bar{s}_n)$ converge dans $\mathbb{R}$, elle y es de Cauchy par le théorème 10.37. Donc il existe un N tel que $p, l > N$ implique

$$\|s_p - s_l\| = \bar{s}_p - \bar{s}_l \leqslant \epsilon. \tag{10.101}$$

Cela signifie que (s_n) est une suite de Cauchy et donc convergente. □

Exemple 10.49(Si l'espace n'est pas complet[1])
Dans un espace pas complet, il est possible de construire un série qui converge absolument sans converger au sens usuel.

Nous allons trouver dans $\mathbb{Q}$ une série qui converge simplement vers $\sqrt{2}$ (et donc ne converge pas dans $\mathbb{Q}$) mais absolument vers 4.

La base est que si $A, B \in \mathbb{Q}$ avec $A < B$ il est possible de résoudre

$$\begin{cases} r_1 + r_2 = A & \text{(10.102a)} \\ |r_1| + |r_2| = B & \text{(10.102b)} \end{cases}$$

pour $r_1, r_2 \in \mathbb{Q}$. Ce n'est pas très compliqué : la solution est $r_1 = (A + B)/2$ et $r_2 = (A - B)/2$.

Nous considérons l'espace $\mathbb{Q}$ qui n'est pas complet dans $\mathbb{R}$. Soit une série (a_k) dans $\mathbb{Q}$ qui converge vers $\sqrt{2}$ (convergence dans $\mathbb{R}$) nous nommons (s_k) la suite des ses sommes partielles. Soit aussi la suite (b_k qui converge vers 4 (zéro serait encore plus facile mais bon, juste pour faire un peu de généralité).

Nous supposons que $a_k < b_k$ pour tout k et que les deux suites sont constituées de rationnels positifs. Nous nommons (s_k) et (s'_k) les sommes partielles. En particulier $s_n < s'_n$ et ce sont des suites croissantes.

Nous savons comment trouver $r_1, r_2 \in \mathbb{Q}$ tels que $r_1 + r_2 = s_1$ et $|r_1| + |r_2| = s'_1$. Par récurrence, si nous savons $r_1, \dots, r_k$ tels que

$$\begin{cases} r_1 + \dots + r_k = s_n & \text{(10.103a)} \\ |r_1| + \dots + |r_k| = s'_n & \text{(10.103b)} \end{cases}$$

(avec, soit dit en passant $k = 2n$), alors nous pouvons trouver des rationnels r_{k+1}, r_{k+2} tels que

$$\begin{cases} r_1 + \dots + r_k + r_{k+1} + r_{k+2} = s_{n+1} & \text{(10.104a)} \\ |r_1| + \dots + |r_k| + |r_{k+1}| + |r_{k+2}| = s'_{n+1}, & \text{(10.104b)} \end{cases}$$

16. Un espace vectoriel normé complet. Typiquement $\mathbb{R}$.

en effet il s'agit de résoudre

$$\begin{cases} r_{k+1} + r_{k+2} = s_{n+1} - r_1 - \ldots - r_k = s_{n+1} - s_n > 0 & \text{(10.105a)} \\ |r_{k+1}| + |r_{k+2}| = s'_{n+1} - |r_1| - \ldots - |r_k| = s'_{n+1} - s'_n > 0. & \text{(10.105b)} \end{cases}$$

Cela se résout comme plus haut. Au final nous pouvons construire une suite (r_k) dans $\mathbb{Q}$ telle que

$$\sum_{k=0}^{2n} r_k = s_n \tag{10.106}$$

et

$$\sum_{k=0}^{2n} |r_k| = s'_n. \tag{10.107}$$

△

Remarque 10.50.
Nous savons que sur les espaces vectoriels de dimension finie toutes les normes sont équivalentes (théorème 10.1). La notion de convergence de série ne dépend alors pas du choix de la norme. Il n'en est pas de même sur les espaces de dimension infinie. Une série peut converger pour une norme mais pas pour une autre.

Lorsque nous verrons la convergence de séries, nous verrons que la convergence normale est la convergence absolue pour la norme uniforme.

Lemme 10.51.
Si E et F sont des espaces de Banach[17] *, l'espace $\mathcal{L}(E, F)$ est également de Banach.*

Démonstration. Soit (u_n) une suite de Cauchy dans $\mathcal{L}(E, F)$; si $x \in E$ il existe N tel que si $l, m > N$ alors $\|a_l - a_m\| < \epsilon$, c'est à dire que pour tout $\|x\| = 1$ on a $\|u_l(x) - u_n(x)\| < \epsilon$. Cela signifie que $u_n(x)$ est une suite de Cauchy dans l'espace complet F. Cette suite converge et nous pouvons définir l'application $u\colon E \to F$ par

$$u(x) = \lim_{n\to\infty} u_n(x). \tag{10.108}$$

Il suffit maintenant de prouver que u est linéaire, ce qui est une conséquence directe de la linéarité de la limite :

$$u(\alpha x + \beta y) = \lim_{n\to\infty} \big(\alpha u_n(x) + \beta u_n(y)\big). \tag{10.109}$$

□

Proposition 10.52 (Thème 40)**.**
Soit E un espace de Banach (espace vectoriel normé complet). Si A est un endomorphisme de E satisfaisant $\|A\| < 1$ pour la norme opérateur, alors $(\mathbb{1} - A)$ est inversible et son inverse est donné par

$$(\mathbb{1} - A)^{-1} = \sum_{k=0}^{\infty} A^k. \tag{10.110}$$

Démonstration. Étant donné que la norme opérateur est une norme algébrique (lemme 10.14), nous avons $\|A^k\| \leqslant \|A\|^k$. Par conséquent la série $\|A^k\|$ est majorée par la série géométrique qui converge. Par conséquent $\sum_k A^k$ est une série absolument convergence et donc convergente par la proposition 10.48 et le fait que $\mathcal{L}(E)$ est complet (proposition 10.51).

17. Je crois qu'il ne faut pas que E soit complet.

Montrons à présent que la somme est l'inverse de $\mathbb{1} - A$ en utilisant le produit terme à terme autorisé par la proposition 10.24 :

$$\sum_{k=0}^{n} A^k(\mathbb{1} - A) = \sum_{k=0}^{n}(A^k - A^{k+1}) = \mathbb{1} - A^{n+1}. \tag{10.111}$$

Par conséquent

$$\|\mathbb{1} - \sum_{k=0}^{n} A^k(\mathbb{1} - A)\| = \|A^{n+1}\| \leqslant \|A\|^{n+1} \to 0. \tag{10.112}$$

□

Proposition 10.53.
Si une série converge dans un espace complet, la norme de son terme général converge vers 0.

Démonstration. Soit $\epsilon > 0$. Si une série converge dans un espace complet, sa suite des sommes partielles est de Cauchy. En particulier il existe un N tel que si $n > N$, nous avons $\|s_n - s_{n-1}\| < \epsilon$. Or évidemment $s_- s_{n-1} = a_n$. Donc pour $n > N$ nous avons $\|a_n\| < \epsilon$. □

Proposition 10.54.
Si la série converge alors la somme est associative

Démonstration. Associativité. Supposons que $\sum_k a_k$ et $\sum_k b_k$ convergent tous deux. Alors nous avons pour tout N :

$$\sum_{k=0}^{N}(a_k + b_k) = \sum_{k=0}^{N} a_k + \sum_{k=0}^{N} b_k. \tag{10.113}$$

Mais si deux limites existent alors la somme commute avec la limite. C'est le cas pour la limite $N \to \infty$, donc

$$\lim_{N\to\infty} \sum_{k=1}^{\infty}(a_k + b_k) = \lim_{N\to\infty} \sum_{k=0}^{\infty} a_k + \lim_{N\to\infty} \sum_{k=0}^{\infty} b_k. \tag{10.114}$$

□

10.7 Série réelle

La notion de série formalise le concept de somme infinie. L'absence de certaines propriétés de ces objets (problèmes de commutativité et même d'associativité) incitent à la prudence et montrent à quel point une définition précise est importante.

10.7.1 Critères de convergence absolue

Étant donné le terme général d'une série, il est souvent –dans les cas qui nous intéressent– difficile de déterminer la somme de la série. L'exemple de la série géométrique est particulier, puisqu'on connait une formule pour chaque somme partielle, mais pour l'exemple des séries de Riemann il n'y a aucune formule simple pour un α général. D'où l'intérêt d'avoir des critères de convergence ne nécessitant aucune connaissance de l'éventuelle limite de la série.

Lemme 10.55 (Critère de comparaison)**.**
Soient $\sum_i a_i$ et $\sum_j b_j$ deux séries à termes positifs vérifiant

$$0 \leqslant a_i \leqslant b_i$$

alors

(1) si $\sum_i a_i$ diverge, alors $\sum_j b_j$ diverge,

(2) si $\sum_j b_j$ converge, alors $\sum_i a_i$ converge (absolument).

Proposition 10.56 (Critère d'équivalence[120])**.**
Soient $\sum_i a_i$ et $\sum_j b_j$ deux séries à termes positifs. Supposons l'existence de la limite (éventuellement infinie) suivante

$$\lim_{i\to\infty} \frac{a_i}{b_i} = \alpha \tag{10.115}$$

avec $\alpha \in \mathbb{R} \cup \{+\infty\}$. Alors

(1) si $\alpha \neq 0$ et $\alpha \neq \infty$, alors

$$\sum_i a_i \text{ converge} \iff \sum_j b_j \text{ converge,} \tag{10.116}$$

(2) si $\alpha = 0$ et $\sum_j b_j$ converge, alors $\sum_i a_i$ converge (absolument),

(3) si $\alpha = +\infty$ et $\sum_j b_j$ diverge, alors $\sum_i a_i$ diverge.

Démonstration. (1) Le fait que la suite a_n/b_n converge vers α signifie que tant sa limite supérieure que sa limite inférieure convergent vers α. En particulier la suite $\frac{a_n}{b_n}$ est bornée vers le haut et vers le bas. À partir d'un certain rang N, il existe M tel que

$$\frac{a_n}{b_n} < M \tag{10.117}$$

et il existe m tel que

$$\frac{a_n}{b_n} > m. \tag{10.118}$$

Nous avons donc $a_n < Mb_n$ et $a_n > mb_n$. La série de (a_n) converge donc si et seulement si la série de (b_n) converge.

(2) Si $\alpha = 0$, cela signifie que pour tout ϵ, il existe un rang tel que $\frac{a_n}{b_n} < \epsilon$, et donc tel que $a_n < \epsilon b_k$. La suite de (a_i) converge donc dès que la suite de (b_i) converge.

(3) Pour tout M, il existe un rang dans la suite à partir duquel on a $\frac{a_i}{b_i} > M$, et donc $a_k > Mb_k$. Si la série de (b_k) diverge, la série de (a_k) doit également diverger.

□

Proposition 10.57 (Critère du quotient[131])**.**
Soit $\sum_i a_i$ une série. Supposons l'existence de la limite (éventuellement infinie) suivante

$$\lim_{i\to\infty} \left|\frac{a_{i+1}}{a_i}\right| = L \tag{10.119}$$

avec $L \in \mathbb{R} \cup \{+\infty\}$. Alors

(1) si $L < 1$, la série converge absolument,

(2) si $L > 1$, la série diverge,

(3) si $L = 1$ le critère échoue : il existe des exemple de convergence et des exemples de divergence.

Démonstration. (1) Soit b tel que $L < b < 1$. À partir d'un certain rang K, on a $\left|\frac{a_{i+1}}{a_i}\right| < b$. En particulier,

$$|a_{K+1}| < b|a_K|, \tag{10.120}$$

et pour a_{K+2} nous avons

$$|a_{K+2}| < b|a_{K+1}| < b^2|a_K|. \tag{10.121}$$

Au final,

$$|a_{K+n}| < b^n|a_K|. \tag{10.122}$$

Étant donné que la série $\sum_{n\geqslant K} b^n$ converge (parce que $b < 1$), la queue de suite $\sum_{i\geqslant K} a_i$ converge, et par conséquent la suite au complet converge.

(2) Si $L > 1$, on a
$$|a_K| < |a_{K+1}| < |a_{K+2}| < \dots \tag{10.123}$$
Il est donc impossible que la suite (a_i) converge vers zéro. La série ne peut donc pas converger.

(3) Par exemple la suite harmonique $a_n = \frac{1}{n}$ vérifie $L = 1$, mais la série ne converge pas. Par contre, la suite $a_n = \frac{1}{n^2}$ vérifie aussi le critère avec $L = 1$ tandis que la série $\sum_n \frac{1}{n^2}$ converge. □

Proposition 10.58 (Critère de la racine[120])**.**
Soit $\sum_i a_i$ une série, et considérons
$$\limsup_{i\to\infty} \sqrt[i]{|a_i|} = L$$
avec $L \in \mathbb{R} \cup \{+\infty\}$. Alors

(1) si $L < 1$, la série converge absolument,

(2) si $L > 1$, la série diverge,

(3) si $L = 1$ le critère échoue.

Démonstration. (1) Si $L < 1$, il existe un $r \in]0, 1[$ tel que $|a_n|^{1/n} < r$ pour les grands n. Dans ce cas, $|a_n| < r^n$, et la série converge absolument parce que la série $\sum_n r^n$ converge du fait que $r < 1$.

(2) Si $L > 1$, il existe un $r > 1$ tel que $|a_n|^{1/n} > r > 1$. Cela fait que $|a_n|$ prend des valeurs plus grandes que n pour une infinité de termes. Le terme général a_n ne peut donc pas être une suite convergente. Par conséquent la suite diverge au sens où elle ne converge pas.

□

10.7.2 Critères de convergence simple

Les critères de comparaison, d'équivalence, du quotient et de la racine sont des critères de convergence absolue. Pour conclure à une convergence simple qui n'est pas une convergence absolue, le critère d'Abel sera notre outil principal.

10.7.2.1 Critère d'Abel

Proposition 10.59 (Critère d'Abel)**.**
Soit la série $\sum_i c_i z_i$ avec

(1) (c_i) est une suite réelle décroissante qui tend vers zéro,

(2) (z_i) est une suite dans $\mathbb{C}$ dont la suite des sommes partielles est bornée dans $\mathbb{C}$, c'est à dire qu'il existe un $M > 0$ tel que pour tout n,
$$\left|\sum_{i=1}^{n} z_i\right| \leqslant M. \tag{10.124}$$
Alors la série $\sum_i c_i z_i$ est convergente.

Remarquons que ce critère ne donne pas de convergence absolue.

Corollaire 10.60 (Critère des séries alternées)**.**
Si (a_n) est une suite décroissante à limite nulle, alors la série
$$\sum_{n=0}^{\infty} (-1)^n a_n \tag{10.125}$$
converge simplement.

10.7.3 Quelques séries usuelles

Exemple 10.61(Série harmonique)
La **série harmonique** est

$$\sum_{i=1}^{\infty} \frac{1}{i} \tag{10.126}$$

et diverge (possède une limite $+\infty$). △

Exemple 10.62(Série géométrique)
La **série géométrique** de raison $q \in \mathbb{C}$ est

$$\sum_{i=0}^{\infty} q^i. \tag{10.127}$$

Étudions la somme partielle $S_N = 1+q+q^2+\cdots+q^N$. Nous avons évidemment $S_N - qS_N = 1-q^{N+1}$ et donc

$$S_N = \sum_{n=0}^{N} q^n = \frac{1-q^{N+1}}{1-q}. \tag{10.128}$$

La limite $\lim_{N\to\infty} S_N$ existe si et seulement si $|q| \leqslant 1$ et dans ce cas nous avons

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}. \tag{10.129}$$

La convergence est absolue.

Si la somme commence en $n = 1$ au lieu de $n = 0$ alors

$$\sum_{n=1}^{\infty} q^n = \frac{1}{1-q} - 1 = \frac{q}{1-q}. \tag{10.130}$$

△

Un cas particulier de la formule (10.128) est le calcul de $\sum_{j=1}^{N} q^{-j}$ bien utile lorsque l'on joue avec des nombres binaires (voir l'exemple 32.12). Nous avons

$$\sum_{j=1}^{N} q^{-j} = \sum_{j=0}^{N} q^{-j} - 1 = \frac{1-q^{-N}}{q-1}. \tag{10.131}$$

Exemple 10.63(Série de Riemann)
Pour $\alpha \in \mathbb{R}$, la **série de Riemann**

$$\sum_{i=1}^{\infty} \frac{1}{i^\alpha} \tag{10.132}$$

converge (absolument, puisque réelle et positive) si et seulement si $\alpha > 1$, et diverge sinon. △

Exemple 10.64(Série exponentielle)
La série exponentielle est la série (pour $t \in \mathbb{R}$)

$$\exp(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}. \tag{10.133}$$

Nous montrons qu'elle converge pour tout $t \in \mathbb{R}$. Si $a_k = t^k/k!$ alors $\frac{a_{k+1}}{a_k} = \frac{t}{k}$ dont la limite $k \to \infty$ est zéro (quel que soit t). En vertu du critère du quotient 10.57 la série exponentielle converge (absolument) pour tout $t \in \mathbb{R}$.

Toutes les propriétés de la fonction de t ainsi définie sont dans le théorème 15.55. △

Exemple 10.65(Série arithméticogémétrique[132])
La **suite arithméticogémétrique** est une suite de la forme $u_{n+1} = au_n + b$ avec a et b non nuls. Si elle possède une limite, cette dernière doit résoudre $l = al + n$, et donc être égale à

$$r = \frac{b}{1-a}. \tag{10.134}$$

Il n'est pas très compliqué de trouver le terme général de la suite en fonction de a et de b. Il suffit de considérer la suite $v_n = u_n - r$, et de remarquer que cette suite est géométrique :

$$v_{n+1} = av_n. \tag{10.135}$$

Par conséquent $v_n = a^n v_0$, ce qui donne pour la suite (u_n) la formule

$$u_n = a^n(u_0 - r) + r. \tag{10.136}$$

△

10.7.4 Moyenne de Cesaro

Si $(a_n)_{n\in\mathbb{N}}$ est une suite dans $\mathbb{R}$ ou $\mathbb{C}$, alors sa **moyenne de Cesaro** est la limite (si elle existe) de la suite

$$c_n = \frac{1}{n}\sum_{k=1}^{n} a_k. \tag{10.137}$$

En un mot, c'est la limite des moyennes partielles.

Lemme 10.66.
Si la suite (a_n) converge vers la limite ℓ alors la suite admet une moyenne de Cesaro qui vaudra ℓ.

Démonstration. Soit $\epsilon > 0$ et $N \in \mathbb{N}$ tel que $|a_n - \ell| < \epsilon$ pour tout $n > N$. En remarquant que

$$\frac{1}{n}\sum_{k=1}^{n} k - \ell = \frac{1}{n}\sum_{k=1}^{n}(a_k - \ell), \tag{10.138}$$

nous avons

$$\begin{aligned}
|\frac{1}{n}\sum_{k=1}^{n} a_k - \ell| &\leqslant |\frac{1}{n}\sum_{k=1}^{N}|a_k - \ell|| + |\frac{1}{n}\sum_{k=N+1}^{n}\underbrace{|a_k - \ell|}_{\leqslant\epsilon}| && \text{(10.139a)}\\
&\leqslant \epsilon + \frac{n-N-1}{n}\epsilon && \text{(10.139b)}\\
&\leqslant 2\epsilon. && \text{(10.139c)}
\end{aligned}$$

Dans ce calcul nous avons redéfinit N de telle sorte que le premier terme soit inférieur à ϵ. □

10.7.5 Écriture décimale d'un nombre

Soit $b \geqslant 2$ un entier qui sera la base dans laquelle nous allons écrire les nombres. Nous considérons l'ensemble $\mathbb{D}_b$ des suites dans $\{0, 1, \ldots, b-1\}$ qui n'ont pas une queue de suite uniquement formée de $b-1$. Autrement dit une suite (c_n) est dans $\mathbb{D}_b$ lorsque pour tout N, il existe $k > N$ tel que $c_k \neq b-1$. Associé à cet ensemble nous considérons la fonction

$$\begin{aligned}
\varphi_b\colon \mathbb{D}_b &\to [0,1[\\
c &\mapsto \sum_{n=1}^{\infty}\frac{c_n}{b^n}.
\end{aligned} \tag{10.140}$$

Lemme 10.67.
La fonction φ_b est bien définie au sens où elle converge et prend ses valeurs dans $[0,1[$.

Démonstration. Tout se base sur la somme de la série géométrique (10.129) sous la forme

$$\sum_{k=0}^{\infty} \frac{1}{b^k} = \frac{b}{b-1}. \tag{10.141}$$

La somme (10.140) est donc majorée par $\sum_n \frac{b-1}{b^n}$ qui converge.

Pour prouver que l'image de φ_b est bien $[0,1[$, nous savons qu'au moins un des c_n (en fait une infinité) est plus petit que $b-1$, donc nous avons la majoration stricte [18]

$$\varphi_b(c) < \sum_{n=1}^{\infty} \frac{b-1}{b^n} = (b-1)\left(\sum_{n=1}^{\infty} \frac{1}{b^n} - 1\right) = 1 \tag{10.142}$$

□

Le fait d'introduire l'ensemble $\mathbb{D}$ au lieu de l'ensemble de toutes les suites est justifié par la proposition suivante. Elle explique pourquoi un nombre possède au maximum deux écritures décimales distinctes et que ces deux sont obligatoirement de la forme, par exemple en base 10 :

$$0.34599999999\ldots = 0.34600000\ldots \tag{10.143}$$

mais qu'un nombre commençant par 0.347 ne peut pas être égal. C'est pour cela que dans la définition de $\mathbb{D}_b$ nous avons exclu les suites qui terminent par tout des $b-1$.

Proposition 10.68.
Soit la fonction

$$\begin{aligned} \varphi\colon \{0,\ldots,b-1\}^{\mathbb{N}} &\to [0,1[\\ x &\mapsto \sum_{n=1}^{\infty} \frac{x_n}{b^n}. \end{aligned} \tag{10.144}$$

Si $\varphi(x) = \varphi(y)$ et si n_0 est le plus petit entier tel que $x_{n_0} \neq y_{n_0}$ alors soit

$$x_{n_0} - y_{n_0} = 1 \tag{10.145}$$

et $x_n = 0$, $y_n = b-1$ pour tout $n > n_0$, soit le contraire : $y_{n_0} - x_{n_0} = 1$ avec $y_n = 0$ et $x_n = b-1$ pour tout $n > n_0$.

Démonstration. Nous nous basons sur la formule (facilement dérivable depuis (10.141)) suivante :

$$\sum_{k=n_0+1}^{\infty} \frac{1}{b^k} = \frac{1}{b^{n_0+1}} \frac{b}{b-1}. \tag{10.146}$$

Nous avons

$$0 = \varphi(x) - \varphi(y) = \frac{x_{n_0} - y_{n_0}}{b^{n_0}} + \sum_{n=n_0+1}^{\infty} \frac{x_n - y_n}{b^n} \geqslant \frac{x_{n_0} - y_{n_0}}{b^{n_0}} - \sum_{n=n_0+1}^{\infty} \frac{b-1}{b^n} = \frac{x_{n_0} - y_{n_0} - 1}{b^{n_0}}. \tag{10.147}$$

Le dernier terme étant manifestement positif [19], il est nul et nous avons $x_{n_0} - y_{n_0} = 1$.

Nous avons donc maintenant

$$0 = \varphi(x) - \varphi(y) = \frac{1}{b^{n_0}} + \sum_{n=n_0+1}^{\infty} \frac{x_n - y_n}{b^n}. \tag{10.148}$$

18. Notez que la somme (10.140) commence à un tandis que la série géométrique (10.141) commence à zéro.

19. C'est ici qu'intervient la subdivision entre le cas $x_{n_0} - y_{n_0} = 1$ ou le contraire. En effet si «ce dernier terme était manifestement *négatif*», il aurait fallu majorer avec de $1-b$ au lieu de $1-b$.

Nous majorons la dernière somme de la façon suivante, en supposant que $|x_n - y_n| \neq b-1$ pour un certain $n > n_0$:

$$\left| \sum_{n=n_0+1}^{\infty} \frac{x_n - y_n}{b^n} \right| \leqslant \sum_{n=n_0+1}^{\infty} \frac{|x_n - y_n|}{b^n} < \sum_{n=n_0+1}^{\infty} \frac{b-1}{b^n} = \frac{1}{b^{n_0}}. \tag{10.149}$$

Étant donné cette inégalité stricte, l'équation (10.148) ne peut pas être correcte (valoir zéro). Nous avons donc $|x_n - b_n| = b-1$ pour tout $n > n_0$. Donc pour chaque $n > n_0$ nous avons soit $x_n = 0$ et $y_n = b-1$, soit $a_n = b-1$ et $b_n = 0$. Pour conclure il faut encore prouver que le choix doit être le même pour tout n.

Nous nous mettons dans le cas $x_{n_0} - y_{n_0} = 1$; dans ce cas nous avons bien l'égalité (10.148) sans petites nuances de signes. Nous écrivons

$$\sum_{n=n_0+1}^{\infty} \frac{x_n - y_n}{b^n} = (b-1) \sum_{n=n_0+1}^{\infty} \frac{(-1)^{s_n}}{b^n} \tag{10.150}$$

où s_n est pair ou impair suivant que $x_n = 0$, $y_n = b-1$ ou le contraire. Si un des $(-1)^{s_n}$ est pas -1 alors nous avons l'inégalité stricte

$$(b-1) \sum_{n=n_0+1}^{\infty} \frac{(-1)^{s_n}}{b^n} > (b-1) \sum_{n=n_0+1}^{\infty} \frac{-1}{b^n} = -\frac{1}{b^{n_0}}. \tag{10.151}$$

Dans ce cas il est impossible d'avoir $\varphi(x) - \varphi(y) = 0$. Nous en concluons que $(-1)^{s_n}$ est toujours -1, c'est à dire $x_n - y_n = 1 - b$, ce qui laisse comme seule possibilité $x_n = 0$ et $y_n = b-1$. □

Théorème 10.69.
L'application $\varphi_b \colon \mathbb{D}_b \to [0,1[$ *est bijective.*

Démonstration. En ce qui concerne l'injection, nous savons de la proposition 10.68 que si $\varphi_b(x) = \varphi_b(y)$ pour $x, y \in \{0, \ldots, b-1\}^{\mathbb{N}}$, alors soit x soit y a une queue de suite composée uniquement de $b-1$, ce qui est exclu dans $\mathbb{D}_b$. Nous en déduisons que φ_b est bien injective en prenant $\mathbb{D}_b$ comme ensemble départ.

La partie lourde est la surjectivité. Nous prenons $x \in [0,1[$ et nous allons construire par récurrence une suite $a \in \mathbb{D}_b$ telle que $\varphi_b(a) = x$. Si il existe $a_1 \in \{0, \ldots, b-1\}$ tel que $x = a_1/b$ alors nous prenons la suite $(a_1, 0, \ldots,)$ et nous avons évidemment $\varphi(a) = x$. Sinon il existe $a_1 \in \{0, \ldots, b-1\}$ tel que

$$\frac{a_1}{b} < x < \frac{a_1 + 1}{b} \tag{10.152}$$

parce que les autres possibilités pour x sont dans l'ensemble $[0,1[\backslash\{\frac{k}{b}\}_{k=0,\ldots,b-1}$ que nous subdivisons en

$$]0, \frac{1}{b}[\cup]\frac{1}{b}, \frac{2}{b}[\cup \ldots \cup]\frac{b-1}{b}, 1[. \tag{10.153}$$

Pour la récurrence nous supposons avoir trouvé $a_1, \ldots, a_n$ tels que

$$\sum_{k=1}^{n} \frac{a_k}{b^k} < x < \sum_{k=1}^{n-1} \frac{a_k}{b^k} + \frac{a_n + 1}{b^n}. \tag{10.154}$$

Encore une fois s'il existe $a_{n+1} \in \{0, \ldots, b-1\}$ tel que $\sum_{k=1}^{n+1} \frac{a_k}{b^k} = x$ alors nous prenons ce a_{n+1} et nous complétons la suite avec des zéros pour avoir $\varphi(a) = x$. Sinon , pour simplifier les notations nous notons $x' = x - \sum_{k=1}^{n} \frac{a_k}{b^k}$ et nous avons

$$0 < x' < \frac{a_n + 1}{b^n}. \tag{10.155}$$

Le nombre x' est forcément dans un des intervalles

$$]\frac{s}{b^{n+1}}, \frac{s+1}{b^{n+1}}[\tag{10.156}$$

avec $s \in \{0, \dots, b-1\}$. Nous prenons le s correspondant à x' comme a_{n+1}. Dans ce cas nous avons

$$\sum_{k=1}^{n+1} \frac{a_k}{b^k} < x < \sum_{k=1}^{n+1} \frac{a_k}{b^k} + \frac{1}{b^{n+1}}. \tag{10.157}$$

Note : les deux inégalités sont strictes. La première parce que s'il y avait égalité, nous nous serions déjà arrêté en complétant avec des zéros. La seconde parce que

$$\sum_{k=n+2}^{\infty} \frac{a_k}{b^k} \leqslant \sum_{k=n+2}^{\infty} \frac{b-1}{b^k} = \frac{1}{b^{n+1}} \tag{10.158}$$

où l'égalité n'est possible que si $a_k = b-1$ pour tout $k \geqslant n+2$. Dans ce cas nous aurions eu

$$x = \sum_{k=1}^{n} \frac{a_k}{b^k} + \frac{a_{n+1}+1}{b^{n+1}} \tag{10.159}$$

et nous aurions choisit le nombre a_{n+1} autrement et complété la suite par des zéros à partir de là. Notons que cela prouve au passage que la suite que nous sommes en train de construire est bien dans $\mathbb{D}_b$ parce qu'elle ne contiendra pas de queue de suite composée de $b-1$.

Ceci termine la construction par récurrence de la suite $a \in \mathbb{D}_b$. Par construction nous avons pour tout $N \geqslant 1$,

$$\sum_{k=1}^{N} \frac{a_k}{b^k} \leqslant x \leqslant \sum_{k=1}^{N} \frac{a_k}{b^k} + \frac{1}{b^{N+1}}, \tag{10.160}$$

autrement dit : $\varphi_b(a_1, \dots, a_N) \in B(x, \frac{1}{b^{N+1}})$. Nous avons donc bien convergence

$$\lim_{N\to\infty} \varphi_b(a_1, \dots, a_N) = x \tag{10.161}$$

et l'application φ_b est surjective. □

L'application $\varphi_b^{-1}\colon [0,1[\to \mathbb{D}_b$ est la **décomposition décimale** en base b des nombres de $[0,1[$.

Tout cela nous permet de montrer entre autres que $\mathbb{R}$ n'est pas dénombrable. Vu qu'il y a une bijection entre $[0,1[$ et $\mathbb{D}_b$, il suffit de prouver que $\mathbb{D}_b$ est non dénombrable. De plus il suffit de démontrer que $\mathbb{D}_b$ est non dénombrable pour un entier $b \geqslant 2$ donné.

Proposition 10.70.
Il n'existe pas de surjection $\mathbb{N} \to \mathbb{D}_b$. Autrement dit $\mathbb{D}_b$ est non dénombrable.

Démonstration. Nous prenons $b \neq 2$ pour des raisons qui seront claires plus tard. Soit $f\colon \mathbb{N} \to \mathbb{D}_b$. Pour $i \in \mathbb{N}$ nous notons

$$f(n) = (c_i^{(n)})_{i\geqslant 1}, \tag{10.162}$$

et nous définissons la suite

$$c_k = \begin{cases} 0 & \text{si } c_k^{(k)} \neq 0 \\ 1 & \text{si } c_k^{(k)} = 0. \end{cases} \tag{10.163}$$

Cela est une suite dans $\mathbb{D}_b$ parce que $b \neq 2$ et que la suite ne contient que des 0 et des 1. Mais nous n'avons $f(n) = c$ pour aucun $n \in \mathbb{N}$ parce que nous avons $c_n \neq f(n)_n$.

Si $b = 2$ alors nous savons que $\mathbb{D}_2 \sim [0,1[\sim \mathbb{D}_3$. Donc $\mathbb{D}_2 \sim \mathbb{D}_3$ et $\mathbb{D}_2$ ne peut pas plus être mis en bijection avec $\mathbb{N}$ que $\mathbb{D}_3$. □

Remarque 10.71.
La preuve ne fonctionne pas en base $b = 2$ parce que rien n'empêche d'avoir une queue de 1. Il y a alors toutefois moyen de se débrouiller en construisant la suite c de façon plus subtile. Si $b = 2$ et $n \in \mathbb{N}$ alors $f(n)$ est une suite de 0 et 1 contenant une infinité de 0 (parce qu'il n'y a pas de queue de suite ne contenant que des 1). Nous construisons alors c de la façon suivante : d'abord

nous recopions $f(0)$ jusqu'à son *deuxième* zéro que nous changeons en 1 ; nommons n_0 le rang de ce deuxième zéro. Ensuite nous recopions les éléments de $f(1)$ à partir du rang n_0+1 jusqu'au second zéro que nous changeons en 1, etc.

Le fait de prendre le deuxième zéro nous garanti que la suite c n'aura pas de queue de suite ne contenant que des 1.

Notons que cette construction s'adapte à tout b ; il suffit de prendre le second terme qui n'est pas $b-1$ et le remplacer par $b-1$.

Corollaire 10.72.
L'ensemble $[0,1[$ *n'est pas dénombrable.*

Démonstration. L'ensemble $[0,1[$ est en bijection avec $\mathbb{D}_b$ que nous venons de prouver n'être pas dénombrable. □

10.8 Produit fini d'espaces vectoriels normés

Dans cette sections nous parlons de produits finis d'espaces. Cela ne signifie pas que chacun des espaces soient séparément de dimension finie.

10.8.1 Norme

La définition de la norme sur un produit d'espaces vectoriels normés découle immédiatement de la définition de la distance 6.161 :

Définition 10.73.
Soient V et W deux espaces vectoriels normés. On appelle ***espace produit*** *de V et W le produit cartésien $V\times W$*

$$V\times W=\{(v,w)\,|\,v\in V,\,w\in W\},\tag{10.164}$$

muni de la norme $\|\cdot\|_{V\times W}$

$$\|(v,w)\|_{V\times W}=\max\{\|v\|_V,\|w\|_W\}.\tag{10.165}$$

Il est presque immédiat de vérifier que le produit cartésien $V\times W$ est un espace vectoriel pour les opération de somme et multiplication par les scalaires définies composante par composante. C'est à dire, si (v_1,w_1), (v_2,w_2) sont dans $V\times W$ et a, b sont des scalaires, alors

$$a(v_1,w_1)+b(v_2,w_2)=(av_1,aw_1)+(bv_2,bw_2)=(av_1+bv_2,aw_1+bw_2).\tag{10.166}$$

Lemme 10.74.
L'opération $\|\cdot\|_{V\times W}\colon V\times W\to\mathbb{R}$ est une norme.

Démonstration. On doit vérifier les trois conditions de la définition 6.31.

- Soit (v,w) dans $V\times W$ tel que $\|(v,w)\|_{V\times W}=\max\{\|v\|_V,\|w\|_W\}=0$. Alors $\|v\|_V=0$ et $\|w\|_W=0$, donc $v=0_V$ et $w=0_W$. Cela implique $(v,w)=(0_v,0_w)=0_{V\times W}$.
- Pour tout a dans $\mathbb{R}$ et (v,w) dans $V\times W$, la norme $\|a(v,w)\|_{V\times W}$ est donnée par $\max\{\|av\|_V,\|aw\|_W\}$. On peut factoriser $\|av\|_V=|a|\|v\|_V$ et $\|aw\|_W=|a|\|w\|_W$ et donc $\|a(v,w)\|_{V\times W}=|a|\max\{\|v\|_V,\|w\|_W\}=|a|\|(v,w)\|_{V\times W}$.
- Soient (v_1,w_1) et (v_2,w_2) dans $V\times W$.

$$\begin{aligned}\|(v_1,w_1)+(v_2,w_2)\|_{V\times W}&=\max\{\|v_1+v_2\|_V,\|w_1+w_2\|_W\}\\&\leqslant\max\{\|v_1\|_V+\|v_2\|_V,\|w_1\|_W+\|w_2\|_W\}\\&\leqslant\max\{\|v_1\|_V,\|w_1\|_W\}+\max\{\|v_2\|_V,\|w_2\|_W\}\\&=\|(v_1,w_1)\|_{V\times W}+\|(v_2,w_2)\|_{V\times W}.\end{aligned}\tag{10.167}$$

□

Toutes ces définitions se généralisent à un produit fini d'espaces vectoriels normés. Si les espaces V_i sont des espaces vectoriels normés, nous pouvons mettre sur le produit une topologie et une norme :

— La topologie produit donnée en 6.5
— La norme maximum $\|v_1, \dots, v_n\|_{max} = \max\{\|v_1\|, \dots, \|v_n\|\}$. Dans le membre de droites, toutes les normes sont différentes.

Une question qui vient est la compatibilité entre ces deux constructions. Est-ce que la topologie associée à la norme maximum est le topologie produit ? Oui.

Lemme 10.75.
La topologie de la norme maximum est la topologie produit[20].

En particulier, pour la topologie de la norme maximum, la convergence d'une suite implique la convergence «composante par composante» par la proposition 6.6.

Proposition 10.76 ([133])**.**
Soient des espaces vectoriels normés V et W ainsi qu'une forme sesquilinéaire $p\colon V \times W \to \mathbb{C}$. Il y a équivalence des faits suivants.

(1) ϕ est continue.
(2) ϕ est continue en $(0,0)$
(3) ϕ est bornée
(4) Il existe $C \geqslant 0$ telle que $|\phi(x,y)| \leqslant C\|x\|\|y\|$ pour tout $(x,y) \in V \times W$.

De plus la norme de ϕ est alors donnée par

$$\|\phi\| = \min\{C \geqslant 0 \text{ tel que } |\phi(x,y)| \leqslant C\|x\|\|y\| \forall (x,y) \in V \times W\}. \tag{10.168}$$

On remarque tout de suite que la norme $\|.\|_\infty$ sur $\mathbb{R}^2$ est la norme de l'espace produit $\mathbb{R} \times \mathbb{R}$. En outre cette définition nous permet de trouver plusieurs nouvelles normes dans les espaces $\mathbb{R}^p$. Par exemple, si nous écrivons $\mathbb{R}^4$ comme $\mathbb{R}^2 \times \mathbb{R}^2$ on peut munir $\mathbb{R}^4$ de la norme produit

$$\|(x_1, x_2, x_3, x_4)\|_{\infty,2} = \max\{\|(x_1, x_2)\|_\infty, \|(x_3, x_4)\|_2\}.$$

Les applications de projection de l'espace produit $V \times W$ vers les espaces «facteurs», V W sont notées $\mathtt{proj}_V$ et $\mathtt{proj}_W$ et sont définies par

$$\begin{aligned} \mathtt{proj}_V\colon V \times W &\to V \\ (v,w) &\mapsto v \end{aligned} \tag{10.169}$$

et

$$\begin{aligned} \mathtt{proj}_W\colon V \times W &\to W \\ (v,w) &\mapsto w. \end{aligned} \tag{10.170}$$

Les inégalités suivantes sont évidentes

$$\begin{aligned} \|\mathtt{proj}_V(v,w)\|_V &\leqslant \|(v,w)\|_{V\times W} \\ \|\mathtt{proj}_W(v,w)\|_W &\leqslant \|(v,w)\|_{V\times W}. \end{aligned} \tag{10.171}$$

La topologie de l'espace produit est induite par les topologies des espaces «facteurs». La construction est faite en deux passages : d'abord nous disons que une partie $A \times B$ de $V \times W$ est ouverte si A et B sont des parties ouvertes de V et de W respectivement. Ensuite nous définissons que une partie quelconque de $V \times W$ est ouverte si elle est une intersection finie ou une réunion de parties ouvertes de $V \times W$ de la forme $A \times B$.

Ce choix de topologie donne deux propriétés utiles de l'espace produit

20. Définition 6.5.

(1) Les projections sont des **applications ouvertes**. Cela veut dire que l'image par $\texttt{proj}_V$ (respectivement $\texttt{proj}_W$) de toute partie ouverte de $V \times W$ est une partie ouverte de V (respectivement W).

(2) Pour toute partir A de V et B de W, nous avons $\text{Int}(A \times B) = \text{Int}\, A \times \text{Int}\, B$.

Une propriété moins facile a prouver est que pour toute partie A de V et B de W nous avons $\overline{A \times B} = \bar{A} \times \bar{B}$. Voir le lemme 10.79.

Ce que nous avons dit jusqu'ici est valable pour tout produit d'un nombre fini d'espaces vectoriels normés. En particulier, pour tout $m > 0$ l'espace $\mathbb{R}^m$ peut être considéré comme le produit de m copies de $\mathbb{R}$.

Exemple 10.77
Si V et W sont deux espaces vectoriels, nous pouvons considérer le produit $E = V \times W$. Les projections $\texttt{proj}_V$ et $\texttt{proj}_W$, définies dans la section 10.8, sont des applications linéaires.

En effet, la projection $\texttt{proj}_V : V \times W \to V$ est donnée par $\texttt{proj}_V(v, w) = v$. Alors,

$$\begin{aligned} \texttt{proj}_V\big((v, w) + (v', w')\big) &= \texttt{proj}_V\big((v + v'), (w + w')\big) \\ &= v + v' \\ &= \texttt{proj}_V(v, w) + \texttt{proj}_V(v', w'), \end{aligned} \tag{10.172}$$

et

$$\texttt{proj}_V\big(\lambda(v, w)\big) = \texttt{proj}_V\big((\lambda v, \lambda w)\big) = \lambda v = \lambda\, \texttt{proj}_V(v, w). \tag{10.173}$$

Nous laissons en exercice le soin d'adapter ces calculs pour montrer que $\texttt{proj}_W$ est également une projection. △

Proposition 10.78.
Si $\mathcal{O}$ est un voisinage de (a, b) dans $V \times W$ alors $\mathcal{O}$ contient un ouvert de la forme $B(a, r) \times B(b, r)$.

Démonstration. Vu que $\mathcal{O}$ est un voisinage, il contient un ouvert et donc une boule

$$B\big((a, b), r\big) = \{(v, w) \in V \times W \text{ tel que } \max\{\|v - a\|, \|w - b\|\} < r\}. \tag{10.174}$$

Évidemment l'ensemble $B(a, r) \times B(b, r)$ est dedans. □

10.8.2 Suites

Nous allons maintenant parler de suites dans $V \times W$. Nous noterons (v_n, w_n) la suite dans $V \times W$ dont l'élément numéro n est le couple (v_n, w_n) avec $v_n \in V$ et $w_n \in W$. La notions de convergence de suite découle de la définition de la norme via la définition usuelle 10.33. Il se fait que dans le cas des produits d'espaces, la convergence d'une suite est équivalente à la convergence des composantes. Plus précisément, nous avons le lemme suivant.

Lemme 10.79.
La suite (v_n, w_n) converge vers (v, w) dans $V \times W$ si et seulement les suites (v_n) et (w_n) convergent séparément vers v et w respectivement dans V et W.

Démonstration. Pour le sens direct, nous devons étudier le comportement de la norme de $(v_n, w_n) - (v, w)$ lorsque n devient grand. En vertu de la définition de la norme dans $V \times W$ nous avons

$$\big\|(v_n, w_n) - (v, w)\big\|_{V \times W} = \max\big\{\|v_n - v\|_V, \|w_n - w\|_W\big\}. \tag{10.175}$$

Soit $\varepsilon > 0$. Par définition de la convergence de la suite (v_n, w_n), il existe un $N \in \mathbb{N}$ tel que $n > N$ implique

$$\max\big\{\|v_n - v\|_V, \|w_n - w\|_W\big\} < \varepsilon, \tag{10.176}$$

et donc en particulier les deux inéquations

$$\|v_n - v\| < \varepsilon \tag{10.177a}$$
$$\|w_n - w\| < \varepsilon. \tag{10.177b}$$

De la première, il ressort que $(v_n) \to v$, et de la seconde que $(w_n) \to w$.

Pour le sens inverse, nous avons pour tout ε un N_1 tel que $\|v_n - v\|_V \leqslant \varepsilon$ pour tout $n > N_1$ et un N_2 tel que $\|w_n - w\|_W \leqslant \varepsilon$ pour tout $n > N_2$. Si nous posons $N = \max\{N_1, N_2\}$ nous avons les deux inégalités simultanément, et donc

$$\max\left\{\|v_n - v\|_V, \|w_n - w\|_W\right\} < \varepsilon, \tag{10.178}$$

ce qui signifie que la suite (v_n, w_n) converge vers (v, w) dans $V \times W$. □

Remarque 10.80.
Il faut remarquer que la norme (10.165) est une norme *par défaut*. C'est la norme qu'on met quand on ne sait pas quoi mettre. Or il y a au moins un cas d'espace produit dans lequel on sait très bien quelle norme prendre : les espaces $\mathbb{R}^m$. La norme qu'on met sur $\mathbb{R}^2$ est

$$\|(x, y)\| = \sqrt{x^2 + y^2}, \tag{10.179}$$

et non la norme «par défaut» de $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ qui serait

$$\|(x, y)\| = \max\{|x|, |y|\}. \tag{10.180}$$

Les théorèmes que nous avons donc démontré à propos de $V \times W$ ne sont donc pas immédiatement applicables au cas de $\mathbb{R}^2$.

Cette remarque est valables pour tous les espaces $\mathbb{R}^m$. À moins de mention contraire explicite, nous ne considérons jamais la norme par défaut (10.165) sur un espace $\mathbb{R}^m$.

Étant donné la remarque 10.80, nous ne savons pas comment calculer par exemple la fermeture du produit d'intervalle $]0, 1, [\times [4, 5[$. Il se fait que, dans $\mathbb{R}^m$, les fermetures de produits sont quand même les produits de fermetures.

Proposition 10.81.
Soit $A \subset \mathbb{R}^m$ et $B \subset \mathbb{R}^m$. Alors dans $\mathbb{R}^{m+n}$ nous avons $\overline{A \times B} = \bar{A} \times \bar{B}$.

La démonstration risque d'être longue ; nous ne la faisons pas ici.

10.8.3 Continuité du produit de matrices

Nous avons introduit des normes sur $\mathbb{M}(n, \mathbb{K})$, entre autres la norme opérateur de la définition 10.5. Qui dit norme dit topologie. Il advient alors la question évidente : est-ce que des opérations aussi élémentaires que le produit de matrices sont continues pour ces topologies ?

Une façon simple de répondre à cela est d'introduire sur $\mathbb{M}(n, \mathbb{K})$ une nouvelle norme très simple : celle de $\mathbb{K}^n$. C'est la topologie par composante. Pour cette topologie, il est simple de voir que le produit matriciel est continu parce que les éléments de AB sont des polynôme en les éléments de A B. Ensuite il suffit d'invoquer l'équivalence de toutes les normes (théorème 10.3).

Voyons comment montrer cela de façon plus directe (bien que le raisonnement précédent soit une démonstration qui devrait déjà avoir convaincu les plus sceptiques). La preuve suivante va donc s'amuser à bien préciser les topologies et caractérisations utilisées.

Lemme 10.82.
Si $\|.\|$ est une norme algébrique sur $\mathbb{M}(n, \mathbb{K})$ ($\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$) alors l'application

$$\begin{aligned} p\colon \mathbb{M}(n, \mathbb{K}) \times \mathbb{M}(n, \mathbb{K}) &\to \mathbb{M}(n, \mathbb{K}) \\ (A, B) &\mapsto AB \end{aligned} \tag{10.181}$$

est continue.

Démonstration. L'espace $\mathbb{M}(n,\mathbb{K}) \times \mathbb{M}(n,\mathbb{K})$ est métrique (définition 10.73), donc la caractérisation séquentielle de la continuité (proposition 6.140) s'applique. Nous considérons donc une suite (A_k, B_k) dans $\mathbb{M}(n,\mathbb{K}) \times \mathbb{M}(n,\mathbb{K})$ convergente vers AB.

Nous savons que la topologie sur $\mathbb{M}(n,\mathbb{K}) \times \mathbb{M}(n,\mathbb{K})$ est la topologie produit (lemme 10.75) et que celle-ci donne la convergence composante par composante dès que nous avons convergence d'une suite ; c'est la proposition 6.6. Nous avons donc $A_k \xrightarrow{\mathbb{M}(n,\mathbb{K})} A$ et $B_k \xrightarrow{\mathbb{M}(n,\mathbb{K})} B$.

Voila pour le contexte. Maintenant, la preuve de la continuité. Nous effectuons les majorations suivantes :

$$\begin{aligned}
\|p(A_k, B_K) - AB\| &\leqslant \|p(A_k, B_k) - p(A_k, B)\| + \|p(A_k, B) - AB\| && \text{(10.182a)}\\
&= \|A_K b_k - A_k B\| + \|A_k B - AB\| && \text{(10.182b)}\\
&= \|A_k(B_k - B)\| + \|(A_k - A)B\| && \text{(10.182c)}\\
&\leqslant \underbrace{\|A_k\|}_{\to\|A\|} \underbrace{\|B_k - B\|}_{\to 0} + \underbrace{\|A_k - A\|}_{\to 0} \|B\|. && \text{(10.182d)}
\end{aligned}$$

□

10.9 Applications multilinéaires

Définition 10.83 (Application multilinéaire)**.**
*Une application $T : \mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_k} \to \mathbb{R}^p$ est dite k-**linéaire** si pour tout $X = (x_1, \ldots, x_k)$ dans $\mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_k}$ les applications $x_i \mapsto T(x_1, \ldots, x_i, \ldots, x_k)$ sont linéaires pour tout i dans $\{1, \ldots, k\}$, c'est à dire*

$$\begin{aligned}
T(\ \cdot\ , x_2, \ldots, x_i, \ldots, x_k) &\in \mathcal{L}(\mathbb{R}^{m_1}, \mathbb{R}^p),\\
T(x_1, \ \cdot\ , \ldots, x_i, \ldots, x_k) &\in \mathcal{L}(\mathbb{R}^{m_2}, \mathbb{R}^p),\\
&\vdots\\
T(x_1, \ldots, x_i, \ldots, x_{k-1}, \ \cdot\ ) &\in \mathcal{L}(\mathbb{R}^{m_k}, \mathbb{R}^p).
\end{aligned} \tag{10.183}$$

*En particulier lorsque $k = 2$, nous parlons d'applications **bilinéaires**. Vous pouvez deviner ce que sont les applications* tri*linéaire ou* quadri*linéaire.*

L'ensemble des applications k-linéaires de $\mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_k}$ dans $\mathbb{R}^p$ est noté $\mathcal{L}(\mathbb{R}^{m_1} \times \ldots \times \mathbb{R}^{m_k}, \mathbb{R}^p)$ ou $\mathcal{L}(\mathbb{R}^{m_1}, \ldots, \mathbb{R}^{m_k}; \mathbb{R}^p)$.

Exemple 10.84
Soit A une matrice avec m lignes et n colonnes. L'application bilinéaire de $\mathbb{R}^m \times \mathbb{R}^n$ dans $\mathbb{R}$ associée à A est définie par

$$T_A(x, y) = x^T A y = \sum_{i,j} a_{i,j} x_i y_j, \qquad \forall x \in \mathbb{R}^m,\ y \in \mathbb{R}^n.$$

△

Nous énonçons la proposition suivante dans le cas d'espaces vectoriels normés [21] parce que nous allons l'utiliser dans ce cas, mais le cas particulier $E_i = \mathbb{R}^{m_i}$ et $F = \mathbb{R}^p$ est important.

Proposition 10.85.
Soient des espaces vectoriels normés E_i et F. Une application n-linéaire

$$T\colon E_1 \times \ldots \times E_n \to F \tag{10.184}$$

est est continue si et seulement s'il existe un réel $L \geqslant 0$ tel que

$$\|T(x_1, \ldots, x_n)\|_F \leqslant L \|x_1\|_{F_1} \cdots \|x_n\|_{F_n}, \qquad \forall x_i \in E_i. \tag{10.185}$$

21. Sans hypothèses sur la dimension.

Démonstration. Pour simplifier l'exposition nous nous limitons au cas $n = 2$ et nous notons $T(x, y) = x * y$

Supposons que l'inégalité (10.185) soit satisfaite.

$$\begin{aligned}\|x * y - x_0 * y_0\| &= \|(x - x_0) * y - x_0 * (y - y_0)\| \\ &\leqslant \|(x - x_0) * y\| + \|x_0 * (y - y_0)\| \\ &\leqslant L\|x - x_0\|\|y\| + L\|x_0\|\|y - y_0\|.\end{aligned} \tag{10.186}$$

Si $x \to x_0$ et $y \to y_0$ on voit que T est continue en passant à la limite aux deux côtes de l'inégalité (10.186).

Soit T continue en $(0, 0)$. Évidemment [22] $0 * 0 = 0$, donc il existe $\delta > 0$ tel que si $x \in B_{E_1}(0, \delta)$ et $y \in B_{E_2}(0, \delta)$ alors $\|x * y\| \leqslant 1$. En particulier si $(x, y) \in B_{E_1 \times E_2}(0, \delta)$ nous sommes dans ce cas. Soient maintenant $x \in E_1\backslash\{0\}$ et $y \in E_2\backslash\{0\}$

$$x * y = \left(\frac{\|x\|}{\delta}\frac{\delta x}{\|x\|}\right) * \left(\frac{\|y\|}{\delta}\frac{\delta y}{\|y\|}\right) = \frac{\|x\|\|y\|}{\delta^2}\left(\frac{\delta x}{\|x\|}\right) * \left(\frac{\delta y}{\|y\|}\right). \tag{10.187}$$

On remarque que $\delta x/\|x\|_m$ est dans la boule de rayon δ centrée en 0_m et que $\delta y/\|y\|_n$ est dans la boule de rayon δ centrée en 0_n. On conclut

$$x * y \leqslant \frac{\|x\|_m\|y\|_n}{\delta^2}.$$

Il faut prendre $L = 1/\delta^2$. □

La norme de T est alors définie comme la plus petite constante L qui fait fonctionner la proposition 10.85.

Définition 10.86.
La norme sur l'espace $\mathcal{L}(E_1 \times \cdots \times E_n, F)$ des applications k-linéaires et continues est

$$\|T\|_{E_1 \times \ldots \times E_n} = \sup\{\|T(u_1, \ldots, u_k)\|_F \,|\, \|u_i\|_{E_i} \leqslant 1, i = 1, \ldots, k\}. \tag{10.188}$$

Nous avons donc automatiquement

$$\|T(u, v)\| \leqslant \|T\|\|u\|\|v\|. \tag{10.189}$$

Et nous notons que cette norme est uniquement définie pour les applications linéaires continues. Ce n'est pas très grave parce qu'alors nous définissons $\|T\| = \infty$ si T n'est pas continue. Cela pour retrouver le principe selon lequel on est continue si et seulement si on est borné.

Proposition 10.87.
On définit les fonctions

$$\begin{array}{rccc}\omega_g : & \mathcal{L}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^p) & \to & \mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)), \\ \omega_d : & \mathcal{L}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^p) & \to & \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^m, \mathbb{R}^p)),\end{array} \tag{10.190}$$

par

$$\omega_g(T)(x) = T(x, \ \cdot\ ), \qquad \forall x \in \mathbb{R}^m,$$

et

$$\omega_d(T)(y) = T(\ \cdot\ , y), \qquad \forall y \in \mathbb{R}^n.$$

Les fonctions ω_g et ω_d sont des isomorphismes qui préservent les normes.

22. Dans la formule suivante, les trois zéros sont les zéros de trois espaces différents.

10.10 Exponentielle de matrice

(1) En ce qui concerne la continuité, nous aurons évidemment besoin de théorie à propos de l'inversion de limites et de sommes. Nous en parlerons donc en 15.12.3.

(2) Les séries entières de matrices seront traitées plus en détail autour de la proposition 15.88.

Proposition 10.88.
Soit V un espace vectoriel de dimension finie et $A \in \mathrm{End}(V)$. La série

$$\exp(A) = \mathbb{1} + A + \frac{A^2}{2} + \frac{A^3}{3} + \ldots = \sum_{k=1}^{\infty} \frac{A^k}{k!}. \tag{10.191}$$

*converge normalement dans $\big(\mathrm{End}(V), \|.\|_{op}\big)$. L'**exponentielle** de la matrice A est cette matrice.*

Démonstration. Vu que la norme opérateur est une norme d'algèbre par le lemme 10.14, nous avons pour tout k la majoration $\|A^k\| \leqslant \|A\|^k$. Nous avons donc

$$\sum_{k=0}^{\infty} \frac{\|A^k\|}{k!} \leqslant \sum_{k} \frac{\|A\|^k}{k!}. \tag{10.192}$$

La dernière somme converge en vertu de la convergence de la série exponentielle donnée en exemple 10.64. □

Étant donné que c'est une limite, il y a une question de convergence et donc de topologie. C'est pour cela que nous ne pouvions pas introduire l'exponentielle de matrice avant d'avoir introduit la norme des matrices. La convergence de la série pour toute matrice sera prouvée au passage dans la proposition 10.89.

La fonction exponentielle $x \mapsto e^x$ n'est pas un polynôme en x, mais nous avons les résultat marrant suivant.

Proposition 10.89.
Si u est un endomorphisme, alors $\exp(u)$ est un polynôme en u[23].

Démonstration. Nous considérons l'application

$$\begin{aligned} \varphi_u \colon \mathbb{K}[X] &\to \mathrm{End}(E) \\ P &\mapsto P(u) \end{aligned} \tag{10.193}$$

Étant donné que l'image de φ_u est un fermé dans $\mathrm{End}(E)$, il suffit de montrer que la série

$$\sum_{k=0}^{\infty} \frac{\varphi_u(X)^k}{k!} \tag{10.194}$$

converge dans $\mathrm{End}(E)$ pour qu'elle converge dans $\mathrm{Image}(\varphi_u)$. Pour ce faire nous nous rappelons de la norme opérateur[24] et de la propriété fondamentale $\|A^k\| \leqslant \|A\|^k$. En notant $A = \varphi_u(X)$,

$$\left\| \sum_{k=n}^{m} \frac{A^k}{k!} \right\| \leqslant \sum_{k=n}^{m} \frac{\|A^k\|}{k!} \leqslant \sum_{k=n}^{m} \frac{\|A\|^k}{k!}, \tag{10.195}$$

ce qui est une morceau du développement de $e^{\|A\|}$. La limite $n \to \infty$ est donc zéro par la convergence de l'exponentielle réelle. La suite des sommes partielles de e^A est donc de Cauchy. La série converge donc parce que nous sommes dans un espace vectoriel réel de dimension finie ($\mathrm{End}(E)$). □

23. Nan, mais j'te jure : exp n'est pas un polynôme, mais $\exp(u)$ est un polynôme de u.
24. Définition 10.5.

Remarque 10.90.
Pourquoi $\exp(u)$ est-il un polynôme d'endomorphisme alors que exp n'est pas un polynôme ? Lorsque nous disons que la fonction $x \mapsto \exp(x)$ n'est pas un polynôme, nous sommes en train de localiser la fonction exp à l'intérieur de l'espace de toutes les fonctions $\mathbb{R} \to \mathbb{R}$, c'est à dire à l'intérieur d'un espace de dimension infinie. Au contraire lorsqu'on parle de $\exp(u)$ et qu'on le compare aux endomorphismes $P(u)$, nous sommes en train de repérer $\exp(u)$ à l'intérieur de l'espace des matrices qui est de dimension finie. Il n'est donc pas étonnant que l'on parvienne moins à faire la distinction.

Si par contre nous considérons exp en tant qu'application $\exp\colon \mathrm{End}(E) \to \mathrm{End}(E)$, ce n'est pas un polynôme.

Si u et v sont des endomorphismes, nous aurons des polynômes P et Q tels que $e^u = P(u)$ et $e^v = Q(v)$; mais nous n'aurons en général évidemment pas $P = Q$. En cela, exp n'est pas un polynôme.

10.11 Mini introduction aux nombres p-adiques

10.11.1 La flèche d'Achille

C'est un grand classique que je donne ici juste comme introduction pour montrer que des série infinies peuvent donner des nombres finis de manière tout à fait intuitive.

Achille tire une flèche vers un arbre situé à 10 m de lui. Disons que la flèche avance à une vitesse constante de 1 m/s. Il est clair que la flèche mettra 10 s pour toucher l'arbre. En 5 s, elle aura parcouru la moitié de son chemin. On le note :

$$\text{temps} = 5s + \dots$$

Reste 5 m à faire. En 2.5 s, elle aura fait la moitié de ce chemin chemin, soit $2.5m = \frac{10}{4}m$. On le note :

$$\text{temps} = \frac{10}{2}s + \frac{10}{4}s +$$

Reste $2.5m$ à faire. La moitié de ce trajet, soit $\frac{10}{8}m$, est parcouru en $\frac{10}{8}s$; on le note encore, mais c'est la dernière fois !

$$\text{temps} = \frac{10}{2}s + \frac{10}{4}s + \frac{10}{8}s +$$

En continuant ainsi à regarder la flèche qui parcours des demi-trajets puis des demi de demi-trajets et encore des demi de demi de demi-trajets, et en sachant que le temps total est $10s$, on trouve :

$$10\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots\right) = 10.$$

On doit donc croire que la somme jusqu'à l'infini des inverse des puissances de deux vaut 1 :

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

Cela peut être démontré à la loyale.

10.11.2 La tortue et Achille

Maintenant qu'on est convaincu que des sommes infinies peuvent représenter des nombres tout à fait normaux, passons à un truc plus marrant.

Achille, qui marche peinard à 10 m/h, part avec $1m$ d'avance sur une tortue qui avance à 1 m/h. Le temps que la tortue arrive au point de départ d'Achille, Achille aura parcouru $10m$, et le temps que la tortue mettra pour arriver à ce point, eh bien, Achille ne sera déjà plus là : il sera à $100m$.

Si la tortue tient bon pendant un temps infini, et si l'on est confiant en le genre de raisonnements faits à la section 10.11.1, elle rattrapera Achille dans

$$1m + 10m + 100m + 1000m + \dots$$

Autant dire que ça ne risque pas d'arriver. Et pourtant, mettons en équations :

$$\begin{cases} x_{\text{Achile}}(t) = 1 + 10t & \text{(10.196a)} \\ x_{\text{tortue}}(t) = t. & \text{(10.196b)} \end{cases}$$

La tortue rejoints Achille au temps t tel que $x_{\text{Achille}(t)} = x_{\text{tortue}}(t)$. Un mini calcul donne $t = -1/9$. Physiquement, c'est une situation logique. Peut-on en déduire une égalité mathématique du style de

$$1 + 10 + 100 + 1000 + \dots = -\frac{1}{9} \text{ ???}$$

Là où les choses deviennent jolies, c'est quand on cherche à voir ce que peut bien être la valeur d'un hypothétique $x = 1 + 10 + 100 + 1000 + \dots$. En effet, logiquement on devrait avoir

$$\begin{aligned} \frac{x}{10} &= \frac{1}{10} + 1 + 10 + 100 + \dots \\ &= \frac{1}{10} + x. \end{aligned}$$

Reste à résoudre l'équation du premier degré : $\frac{x}{10} = x + \frac{1}{10}$. Ai-je besoin de donner la solution ?

10.11.3 Dans les nombres p-adiques, c'est vrai

Nous nous proposons d'apprendre sur les nombres p-adiques juste ce qu'il faut pour montrer que l'égalité

$$\sum_{k=0}^{\infty} 10^k = -\frac{1}{9} \tag{10.197}$$

est vraie dans les nombres 5-adiques. Tout ce qu'il faut est sur wikipedia.

Soit $a \in \mathbb{N}$ et p, un nombre premier. La **valuation** p-adique de a est l'exposant de p dans la décomposition de a en nombres premiers. On la note $v_p(a)$. Pour un rationnel on définit

$$v_p\left(\frac{a}{b}\right) = v_p(a) - v_p(b) \tag{10.198}$$

La **valeur absolue** p-adique de $r \in \mathbb{Q}$ est

$$|r|_p = p^{-v_p(r)}. \tag{10.199}$$

Nous posons $|0|_p = 0$. De là nous considérons la distance

$$d_p(x, y) = |x - y|_p. \tag{10.200}$$

Lemme 10.91.
L'espace $(\mathbb{Q}, d_p)$ est un espace métrique[25].

Nous considérons maintenant $p = 5$. Étant donné que $a = 5 \cdot 2$ nous avons $v_5(10) = 1$ et

$$v_5\left(\frac{1}{9}\right) = v_5(1) - v_5(9) = 0. \tag{10.201}$$

Nous avons

$$\sum_{k=0}^{N} 10^k + \frac{1}{9} = \frac{10^{N+1}}{9} \tag{10.202}$$

25. Définition 6.18

mais

$$v_p\left(\frac{10^{N+1}}{9}\right) = v_5(10^{N+1}) - v_5(9) = N+1. \tag{10.203}$$

Par conséquent

$$d_5\big(\sum_{k=0}^{N} 10^k, -\frac{1}{9}\big) = |\frac{10^{N+1}}{9}|_p = p^{-(N+1)}. \tag{10.204}$$

En passant à la limite,

$$\lim_{N\to\infty} d_5\big(\sum_{k=0}^{N} 10^k, -\frac{1}{9}\big) = 0, \tag{10.205}$$

ce qui signifie que [26]

$$\sum_{k=0}^{\infty} 10^k = -\frac{1}{9}. \tag{10.206}$$

26. Voir la définition 10.43 de la convergence d'une série dans un espace métrique.